AFCRL-67-0499

AD 660595

THE EFFECTS OF TASK-INDUCED STRESS ON SPEECH

Michael H. L. Hecker
Kenneth N. Stevens
Gottfried von Bismarck
Carl E. Williams

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02138

Contract No. AF 19(628)-6052 new
Project No. 5628

FINAL REPORT

Period Covered: 1 July 1966 through 30 June 1967

25 August 1967

Contract Monitor: Philip Lieberman
Data Sciences Laboratory

Distribution of this document is unlimited. It may be released to the Clearinghouse, Department of Commerce, for sale to the general public.

Prepared for:

AIR FORCE CAMBRIDGE RESEARCH LABORATORIES
OFFICE OF AEROSPACE RESEARCH
UNITED STATES AIR FORCE
BEDFORD, MASSACHUSETTS

NOV 1 1967

Reproduced by the
CLEARINGHOUSE
for Federal Scientific & Technical
Information Springfield Va. 22151

81

AFCRL-67-0499

THE EFFECTS OF TASK-INDUCED STRESS ON SPEECH

Michael H. L. Hecker
Kenneth N. Stevens
Gottfried von Bismarck
Carl E. Williams

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02138

Contract No. AF 19(628)-6052
Project No. 5628

FINAL REPORT

Period Covered: 1 July 1966 through 30 June 1967

25 August 1967

Contract Monitor: Philip Lieberman
Data Sciences Laboratory

Distribution of this document is unlimited. It may be released to the Clearinghouse, Department of Commerce, for sale to the general public.

Prepared for:

AIR FORCE CAMBRIDGE RESEARCH LABORATORIES
OFFICE OF AEROSPACE RESEARCH
UNITED STATES AIR FORCE
BEDFORD, MASSACHUSETTS

THE EFFECTS OF TASK-INDUCED STRESS ON SPEECH

Abstract

In order to induce stress in an experimental subject, a task involving the addition of numbers under time pressure was developed. The subject was required to read six meters and to announce the sum of his readings, together with a test word. By controlling the duration of the meter display, the experimenter could vary the level of stress induced in the subject. For each of ten subjects, numerous verbal responses were obtained while the subject was under stress and while he was relaxed.

Contrasting responses containing the same test word were assembled into paired-comparison listening tests. Listeners could identify the stressful responses of some subjects with better than 90 percent accuracy and of others only at chance level. The test words from contrasting responses were analyzed with respect to level and fundamental frequency, and spectrograms of these test words were examined. The results indicate that stress can produce a number of characteristic changes in the acoustic speech signal. Most of these changes are attributable to modifications in the amplitude, frequency, and detailed waveform of the glottal pulses. Other changes result from differences in articulation. Although the manifestations of stress varied considerably from subject to subject, the test words of most subjects exhibited some consistent effects.

TABLE OF CONTENTS

LIST OF CONTRIBUTORS

The following scientists and engineers contributed to the research reported in this document:

Michael H. L. Hecker
Kenneth N. Stevens
Gottfried von Bismarck
Carl E. Williams
Warren M. Brodey
Vincent J. Sharkey

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Modern technology has produced machinery which requires the human operator to perform complex tasks under hazardous conditions. The resulting psychological or physiological stress may reduce his normal proficiency and limit his physical abilities. Numerous studies have been concerned with the effects of various kinds of stress on performance.* Although it is recognized that performance often involves some form of verbal communication, few studies have considered the possibility that speech may serve as an indicator of stress.[1,6,7] If it can be demonstrated that stress is manifested in the acoustic speech signal, it may be possible to develop a technique for analyzing a speaker's voice to determine when he is under stress. Eventually, it may be feasible to build an electronic device which will accomplish such an analysis automatically.

Various definitions of stress have been employed in the published studies. Some experimenters have defined stress in terms of the physical stimulus. Among the arguments advanced in favor of this definition is that the stimulus can be accurately measured and easily described. However, a description of the stimulus provides no information about the behavior of people who are exposed to it. It is generally

---

* These studies are reviewed in References 2 and 5.

known that individuals differ appreciably in their reactions to the same stimulus.[2]

Stress may also be defined in terms of the response to a stimulus. Several physiological and psychological measures have been used as indicators of stress. But these measures often change independently of one another and are influenced by other variables, such as motivation and performance. It is difficult to decide, therefore, which response measures should be emphasized in a given experiment. Another limitation of this definition is that physiological and psychological data are difficult to interpret.

Since stress cannot be satisfactorily defined in terms of either the stimulus or the response, many investigators have treated stress as an intervening variable. They have decided that stress occurs when a particular situation threatens the attainment of some goal.[5] Because this motivation-oriented definition has been found to be practically useful, it was also employed in the present study. Motivated behavior can be best established in a specific task, and this task can be manipulated by the experimenter to frustrate the subject. Inferences about the level of stress induced in the subject can be made from his reactions to the task and from his performance.

The objective of this study was to provide a qualitative answer to the question: How is task-induced stress manifested in the acoustic speech signal? From a knowledge of the process of speech production it is possible, of course, to speculate on what portions of the speech signal are likely to be influenced by stress. The respiratory function and the operation of the larynx are probably less organized in a stressful situation; these effects would principally alter the manner of excitation of the vocal apparatus, especially during the production of voiced sounds. Stress may also influence articulation. But to what extent can such changes be demonstrated experimentally? Are such changes reliable indicators of stress, independent of the speaker and the contents of the utterance? The general approach used in this study was to develop a task which could induce stress in an experimental subject and which would require him to respond verbally. For a number of subjects, the identical speech material was recorded during stress and control conditions, and a detailed acoustical analysis of this material was then undertaken.

2. PRELIMINARY STUDY

A preliminary study was carried out to develop criteria for the design of a suitable task. This study employed the concept of workload.[4] Basic to this concept is the observation that a person can perceive and react to only so much information in a given time interval before a condition of overload is reached which is regarded as stressful. The maximum capacity of a human monitor of instruments can be estimated from motion-picture records of eye fixations on the instruments. For instruments of the type used in aircraft, a set of signals having a total bandwidth of about 1.25 Hz represents a full load for the monitor.[8]

The first effort consisted of implementing a visual monitoring task. To provide the experimental subject with a controlled visual environment, a three-panel cubicle measuring 3 x 3 x 3 feet was constructed. So as not to interfere with the recording of the subject's speech, the panels of the cubicle were made of wooden frames covered with wire mesh and burlap. Six 3-inch round 50-0-50 microampere meters were installed in the center panel; three meters were mounted 9 1/2 inches apart in each of two horizontal rows which were separated by 19 inches. The cubicle was supported by a table which was covered with a layer of fiberglass and topped with perforated masonite. The height of the table was such that when the subject was seated the level of his eyes was midway between the two rows of meters.

The meters were activated by six continuously varying random signals of different bandwidths. These signals were reproduced from a 7-track instrumentation tape. When the tape was reproduced at a speed of 1 7/8 inches per second, the bandwidths of the six signals were: 0.03, 0.05, 0.12, 0.20, 0.32, and 0.48 Hz. Since the total bandwidth was 1.20 Hz, the visual display very nearly provided a full load for the subject. By simply changing the tape speed, the effective bandwidths of the signals could be changed proportionately; the subject could thus be loaded to a lesser or greater degree.

The monitoring task involved the following experimental procedure. Seated before the cubicle, but without observing the meters, the subject first recorded selected speech material onto a Master Tape. The speech material consisted of monosyllabic and polysyllabic words. The material was read from a typed page at a normal rate, allowing at least one second between successive utterances. A control condition was then established in which the subject heard his Master Tape over earphones and responded by repeating back the words he heard. His responses were recorded on a second tape. The meters were not activated during this control condition.

Then followed four experimental conditions. During Condition I, the meters were activated to provide a 100 percent workload. The subject was instructed to push a silent hand switch whenever any of the meter deflections exceeded either +40 (microamperes) in the positive direction or -40 in the negative direction. He also repeated the words from his Master Tape as heard over the earphones, exactly as during the control condition.

Condition II was identical to Condition I except that the instrumentation tape was played at twice the previous speed. This increased the effective bandwidths of the meter-control signals by a factor of two, but the response characteristics of the meter movements were such that the average number of meter deflections above +40 and below -40 was not exactly doubled.

For Conditions III and IV neither the Master Tape, the earphones, nor the hand switch were employed. The instrumentation tape was played at 1 7/8 inches per second for Condition III, providing a 100 percent workload, and at twice this speed for Condition IV. In both conditions, the subject responded by calling out the number of any meter on which he observed a deflection above +40 or below -40, and the direction of the deflection. Thus, a typical response was: "Five, low."

The subjects who participated in these experiments were divided in their opinion as to whether the task actually produced stress. Their solicited comments indicated that the task possessed a number of serious limitations. One problem was that in Conditions I and II the visual display and response (hand switch) and the auditory presentation and response (repetition of utterances heard) were treated as separate processes. This division of the task into two unrelated sub-tasks appeared to be undesirable from the point of view of subject involvement. Conditions I and II did not seem sufficiently realistic or important to some subjects for them to become fully involved in the task.

Another apparent difficulty was that in all experimental conditions the rate of speech could have been affected directly by the task as well as indirectly by the task-induced stress. The subject could have viewed the verbal response as an activity which, since it is likely to interfere with his continuing monitoring job, should be accomplished as rapidly as possible. Considering the exploratory aspect of these early efforts, and the fact that the intended acoustic data reduction would be very time consuming, the recordings obtained in these experiments were not systematically analyzed.

That some subjects did not consider the task stressful may be related in part to inadequate motivation and in part to the finding of other investigators that different individuals have widely different tolerances for stress.[2] To be able to induce stress in a given individual, it appeared necessary that the experimenter be able to vary the level of stress, i.e., the difficulty of the task, while the experiment is progressing. The experimenter would be able to infer how much stress the task is inducing in a subject by observing the subject's reaction to the task and his performance. On the basis of such observations, the experimenter could change the difficulty of the task to find the limit to which the subject can be "pushed" before his performance deteriorates markedly.

Since the difficulty of the monitoring task described above was determined primarily by the speed of the instrumentation tape, which could not be continuously varied, it was decided that this task should be abandoned. The experience gathered during this preliminary study culminated in the development of criteria for the design of an improved task. The new task was to incorporate the following features:

1. The task would provide a measure of performance which is closely related to the subject's verbal activity and which can be conveniently observed by the experimenter.

2. The difficulty of the task would be continuously variable during the experimental session.

3. The task would be structured so that the subject's visual and verbal activities are integrated.

4. The task would appear realistic and meaningful to the subject so as to motivate him and invite his deep involvement.

5. Further motivation would be provided by paying the subject in proportion to his performance.

6. As the experiment progresses, the subject would be periodically informed as to the current success of his efforts.

3. THE ADDITION TASK

A new visual monitoring task was developed. In this task, the subject was required to add a set of meter readings under time pressure. The task employed the previously described cubicle with the six meters, but the meters were not activated from the instrumentation tape. Instead, the meters were controlled by means of attenuator networks which could be programmed to provide different steady-state meter deflections. All meters were energized and de-energized together. The display duration, which determined the difficulty of the task, was continuously variable. Concurrent with the presentation of each display, the subject was given a test word by the experimenter. The subject mentally added the six meter readings, and when the display was terminated he announced his numerical answer, together with the test word.

The correct answer corresponding to each display was determined in advance, so that the experimenter could make inferences about the level of stress induced in the subject by checking the correctness of the announced answers. By varying the display duration, the experimenter could explore a given subject's tolerance for stress in an effort to produce as dramatic a change in his voice as possible. The strategy employed by the experimenter necessarily varied from subject to subject.

3.1 Pilot Experiments

Several pilot experiments were carried out to test the effectiveness of the new task and experimental procedure. The instrumentation used in these experiments was relatively simple: The experimenter programmed a particular meter display by adjusting six decade attenuators and six switches which controlled the deflections and polarities of the six meters. He then energized the attenuators for a certain length of time to present the display to the subject. While the subject was busy with this display, the experimenter programmed a second set of attenuators and switches which would determine the next display. In these and all subsequent experiments, the subject performed the task in a sound-treated studio, while the experimenter was located in an adjoining control room. The experimenter could speak to the subject over a loudspeaker in the studio.

Before a given display was presented, the experimenter read a test word to the subject. The subject was required to repeat this test word, whereupon the experimenter presented the display. During the display period, the subject attempted to read each meter (to the nearest multiple of 5 microamperes) and to determine the algebraic sum of his readings. When the display was terminated, he announced his answer and again said the test word. The experimenter then told the subject whether his answer was correct or incorrect. On the basis of the subject's performance on ten successive displays, the experimenter varied the display time until a performance level of 50 percent was reached. A control condition was established by asking the subject to read only one meter, and by continuing the display until the subject responded.

A number of conclusions were drawn from these pilot experiments. Most of the subjects who participated in these experiments, including the experimenters themselves, found the task quite stressful. They were sufficiently motivated by the intellectual challenge of correct and rapid mental addition to become very involved in the task. It was therefore concluded that the addition task was suitable for the intended purpose.

A few subjects stated that they might have been involved even further if they had been given some control over the display duration. They suggested that the amount of money paid for a correct answer should be inversely related to the display duration; more money should be paid for answers corresponding to shorter durations. Then there would be a trade-off in reward between speed and accuracy, and the subject may wish to work with a display duration that best matches his personal approach to the task. It was decided that the final instructions for the task would reflect this suggestion, and that a signalling system would be provided to enable the subject to request an increase or decrease in display duration.

Another conclusion drawn from the pilot experiments was that the experimenter should not speak with the subject. The subject may unconsciously attempt to duplicate the experimenter's particular pronounciation of the test words, and this might possibly interfere with, or obscure, the effects of stress in the subject's voice. Furthermore, the manner in which the experimenter tells the subject whether his answer is right or wrong represents an uncontrolled variable; the experimenter may unintentionally encourage or discourage the subject. It was therefore decided that the final version of the task would employ selectively illuminated panels which would display the test words in the cubicle. Also, the subject would be informed of his performance by means of light signals.

During the pilot experiments it became evident that only very rough inferences could be made about the level of stress induced in the subject. The experimenter had access to only two sources of information on which he could base his estimate of stress: the subject's performance and his speech. Additional means for estimating the level of stress appeared desirable, so that the experimenter would be in a better position to devise a particular strategy for a given subject. A finer strategy would possibly make the task more challenging and increase the efficiency of the experiment. It was decided that the experimenter should be able to observe the subject's face and monitor one or two physiological measures which would indicate the subject's response to the task.

Some subjects tended to move their heads from side to side as they read the individual meters. Occasionally they also leaned forwards or backwards; this behavior could have been a result of their anxiety. Because of this head motion, the distance between the subject's mouth and the microphone in the cubicle varied considerably. At different times the microphone sampled different portions of the sound field generated by the subject. Considering that changes in the level and spectrum of a speech signal may be important indicators of stress, this situation was considered undesirable. The use of a fixed microphone seemed incongruous with the aims of this study. It was concluded, therefore, that further experiments would employ a small high-quality microphone which would be attached to a headband worn by the subject.

3.2 Instrumentation of Task

A block diagram of the instrumentation used in the final version of the addition task is shown in Fig. 1. The experimenter could observe the subject through a small horizontal window in the studio wall and a matching opening in the center panel of the cubicle. The opening in the cubicle was located between the two rows of meters. When the studio was well lighted but dim work lights were used in the control room, the subject could not see the experimenter. By means of an electric timer and programmable display control circuits, the experimenter could present a rapid sequence of meter displays of various durations. Together with each meter display, one of five test words was displayed by a light box in the cubicle. Light signals were used to inform the subject whether his answer was right or wrong, and also to indicate when the subject wanted the display duration to be increased or reduced.

As a safety precaution, the subject's speech was recorded with two tape recorders. One channel of a two-track tape recorder was used to record comments by the experimenter about the subject's reactions to the task. After rectification and low-pass filtering, the subject's speech was also recorded on one channel of a chart recorder. A strain-gauge respirometer, which was strapped around the subject's chest, provided a measure of his respiratory activity. Strain gauges were installed in the chair in which the subject sat to provide a measure of his movement. Surface electrodes were applied to the forearm and hand of the subject to provide a measure of his skin potential. These three measures were also recorded with the chart recorder. All instructions were pre-recorded and presented to the subject over a loudspeaker.

A photograph of the control room is shown in Fig. 2. The chart recorder was positioned next to the observation window, and a mirror was installed over the window so that the experimenter could conveniently observe both the subject's face and his recorded responses. The electric timer, which is seen on the table to the left of the chart recorded, was operated by an assistant to the experimenter. On the other side of the window are the two tape recorders. This photograph does not show the display control circuits, which were programmed by a second assistant. By using assistants, the experimenter was freed of technical responsibilities and could concentrate fully on the subject. The experimenter and his assistants listened to the subject's speech over earphones.

Figure 3 shows a photograph of a subject watching the displays in the cubicle. The opening through which the subject was observed appeared as a black rectangle between the two rows of meters. Below the bottom row of meters was the light box displaying the test words. The subject wore a headband on which the recording microphone was mounted. The photograph also shows the push buttons which the subject used to request a longer or shorter display duration.

The display control circuits deserve further discussion. A block diagram of these circuits is shown in Fig. 4. Ten different meter displays could be programmed at one time by means of ten plugs. Each plug contained a polarity strapping and a shunt resistor for each of six meter circuits. When the timer was

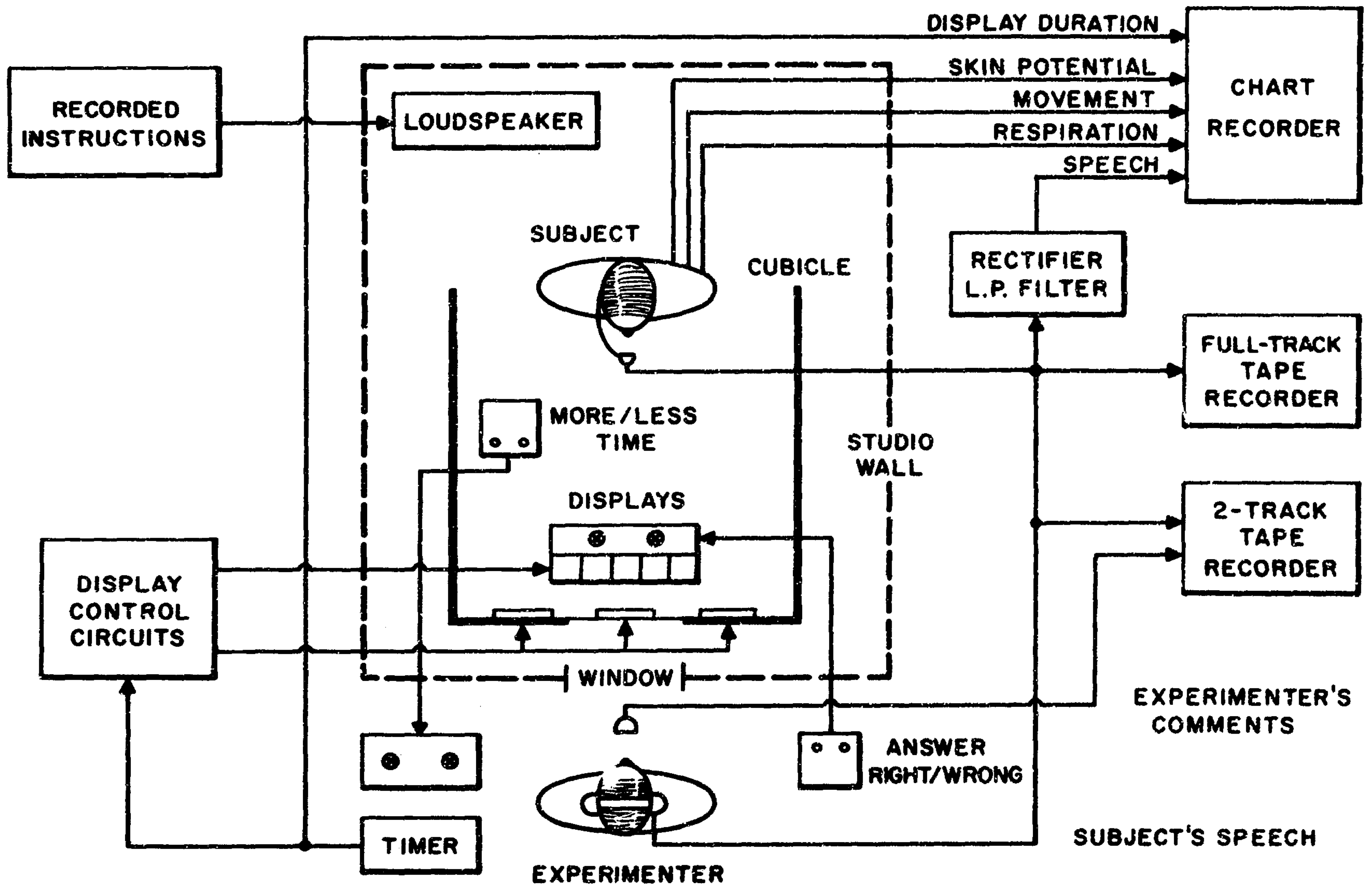


FIG. 1 BLOCK DIAGRAM OF INSTRUMENTATION USED IN ADDITION TASK.

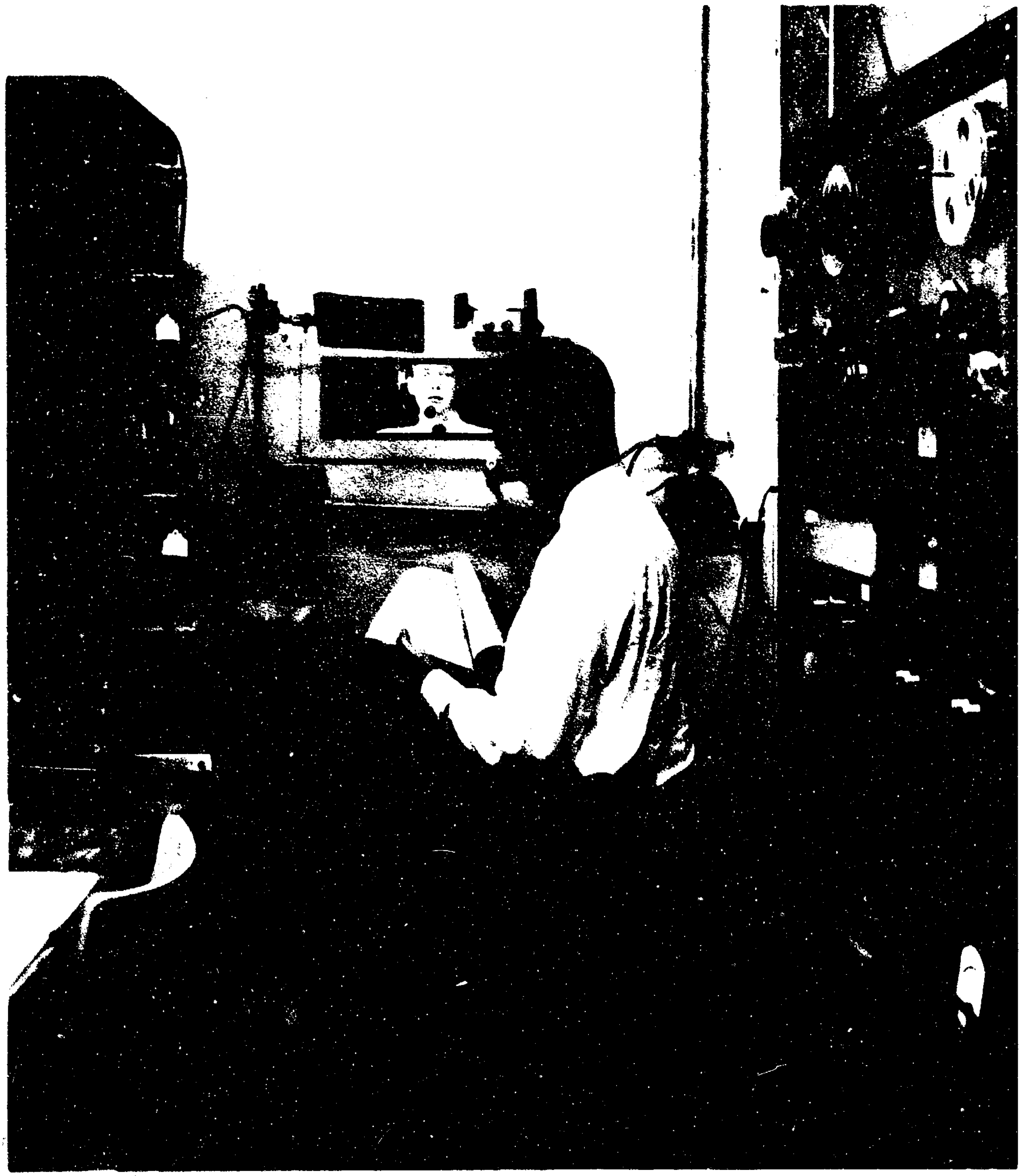


FIG. 2 PHOTOGRAPH OF CONTROL ROOM SHOWING EXPERIMENTER SEATED BY OBSERVATION WINDOW AND CHART RECORDER.

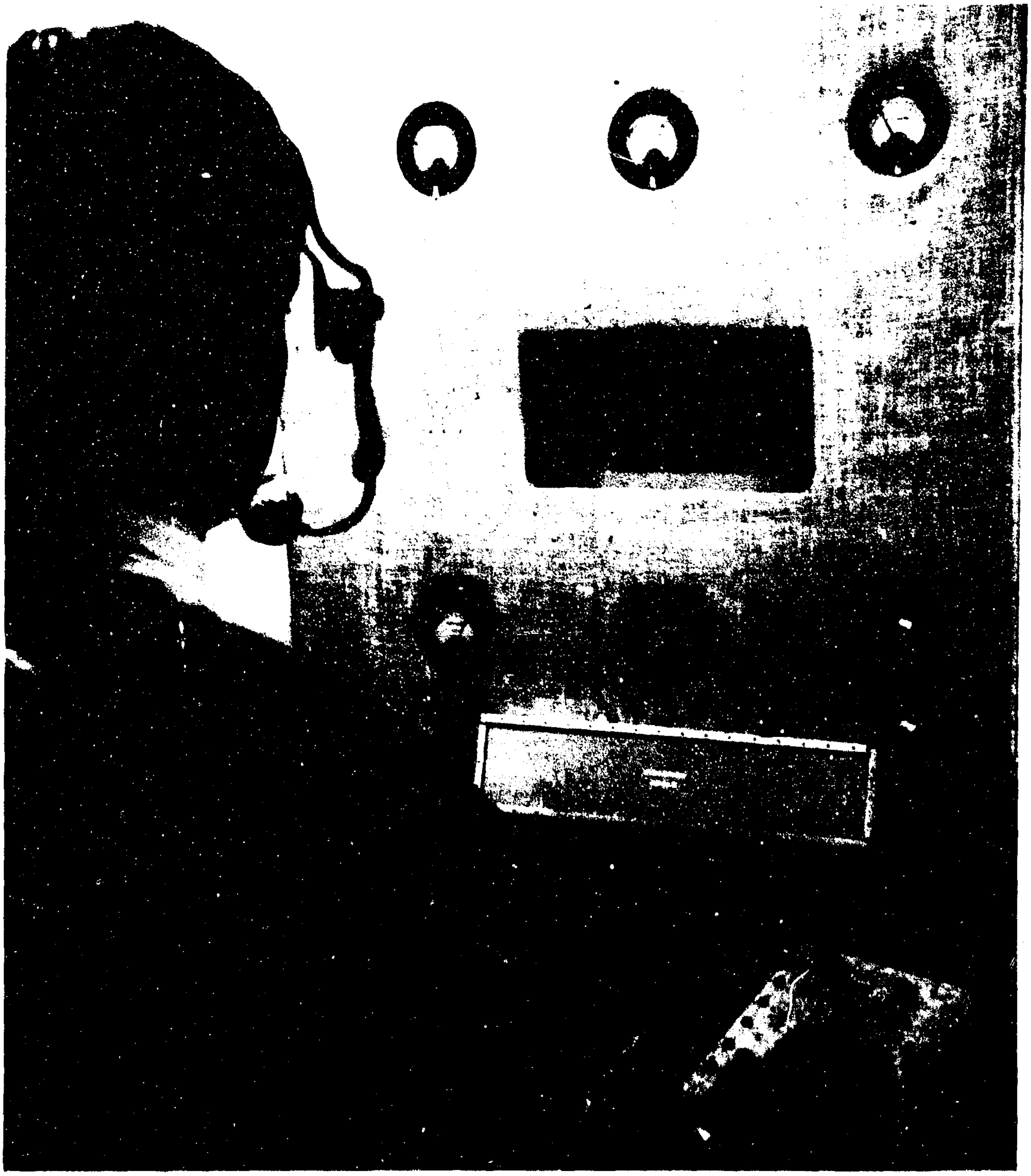


FIG. 3 PHOTOGRAPH OF SUBJECT WATCHING METER AND TEST-WORD DISPLAYS IN CUBICLE.

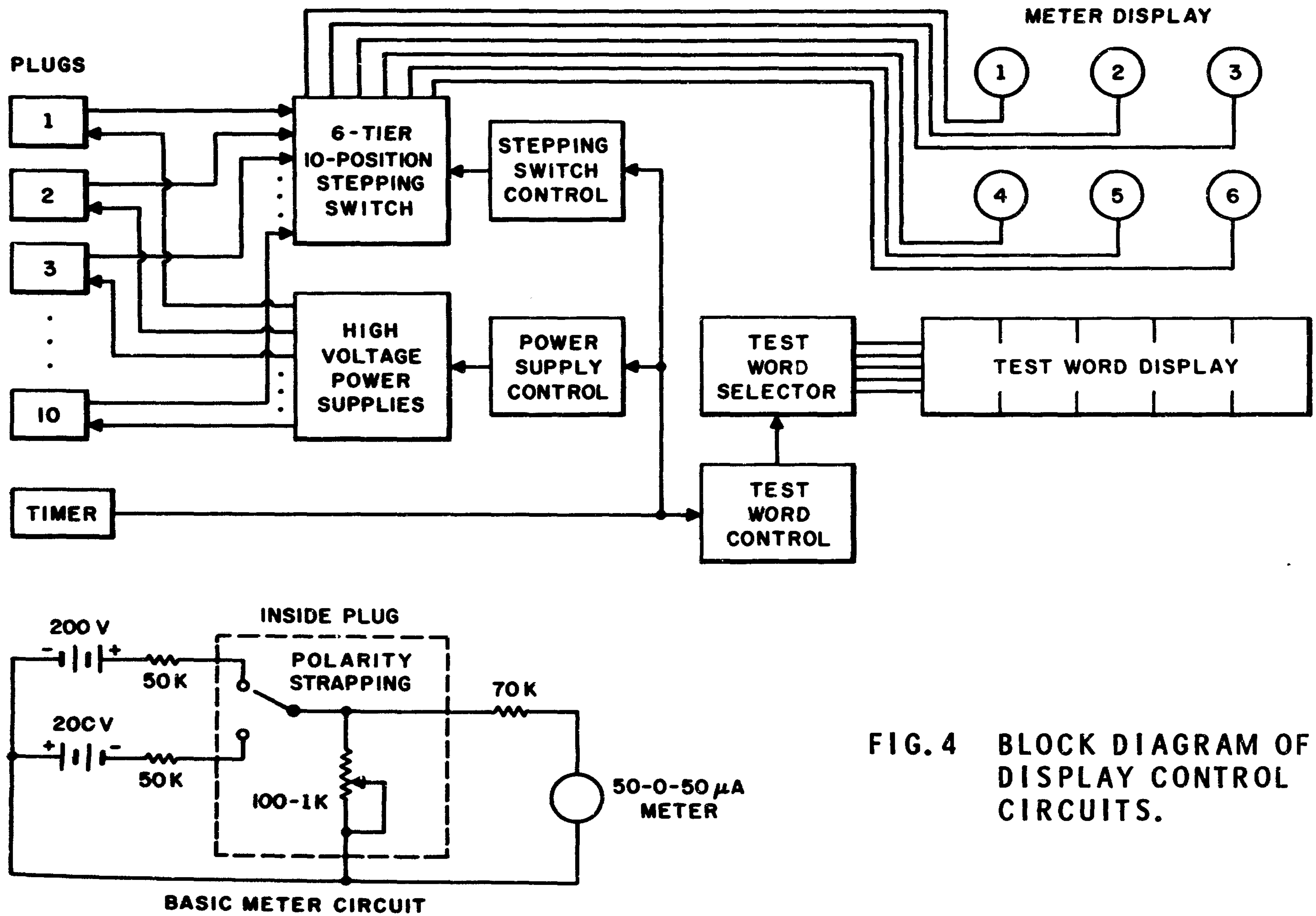


FIG. 4 BLOCK DIAGRAM OF DISPLAY CONTROL CIRCUITS.

operated, two high-voltage power supplies were energized and all plugs were furnished with positive and negative voltages. But only one plug was selected by a stepping switch to control the meter display. When the time interval set on the timer had elapsed, the voltages were removed from all plugs, which terminated the meter display, and the stepping switch automatically advanced the meter connections to the next plug. The timer could then be operated again to present the new meter display.

Lights mounted above the ten sockets holding the plugs indicated which plug was selected by the stepping switch. After several plugs were passed over, the operator replaced these plugs with others so that the experimenter did not have to wait for new meter displays to be programmed. Forty plugs were available, and the operator used these plugs according to a randomization schedule. Because each plug was marked with the correct answer, the operator could signal to the subject and tell the experimenter whether an announced answer was right or wrong.

The timer also controlled the display of the five test words. For the duration of each meter display, one test word was illuminated in the light box. After the termination of a given meter and test-word display, a multiple-position stepping switch selected the next test word. The contacts of this stepping switch were strapped to approximate a random sampling of the five test words. This arrangement insured independence between the meter displays and the test-word displays. Thus, if a given meter display was presented more than once during an experiment, it was likely to be coupled with a different test word each time. There was no provision for informing the operator which particular test word was displayed with each meter display.

The operator did not know, therefore, whether the subject always announced the displayed test words or made occasional mistakes. In this study, confusions among the test words were not treated as indicators of stress.

The test words were, in fact, phrases. Although each phrase contained two words, it will be referred to as a test word in order to provide a uniform terminology throughout this report. The five test words were: "Generator Current," "Oil Pressure," "Capsule Temperature," "Relative Velocity," and "Deviation Angle." These words were chosen primarily because it was thought that they could be naturally associated with meter readings. Earlier studies have demonstrated the importance of making the task appear as realistic and meaningful as possible. The test words also provided a variety of speech sounds which could be examined for evidence of stress. Without prior knowledge of how stress might influence different speech sounds, there were no criteria for selecting test words which would represent particular classes of speech sounds.

Figure 5 shows a portion of a response record obtained with the four-channel chart recorder. This record covers a time interval in which six meter displays were presented to the subject. The meter displays are indicated by the heavy lines in the lower margin of the record. These displays were assigned consecutive numbers, as indicated, in order to facilitate the later selection of verbal responses for analysis. Toward the end of each display, the first tracing shows the subject's verbal response. The tracing labeled respiration shows intervals of irregular and shallow breathing. As would be expected, this tracing changes in a characteristic manner whenever the subject

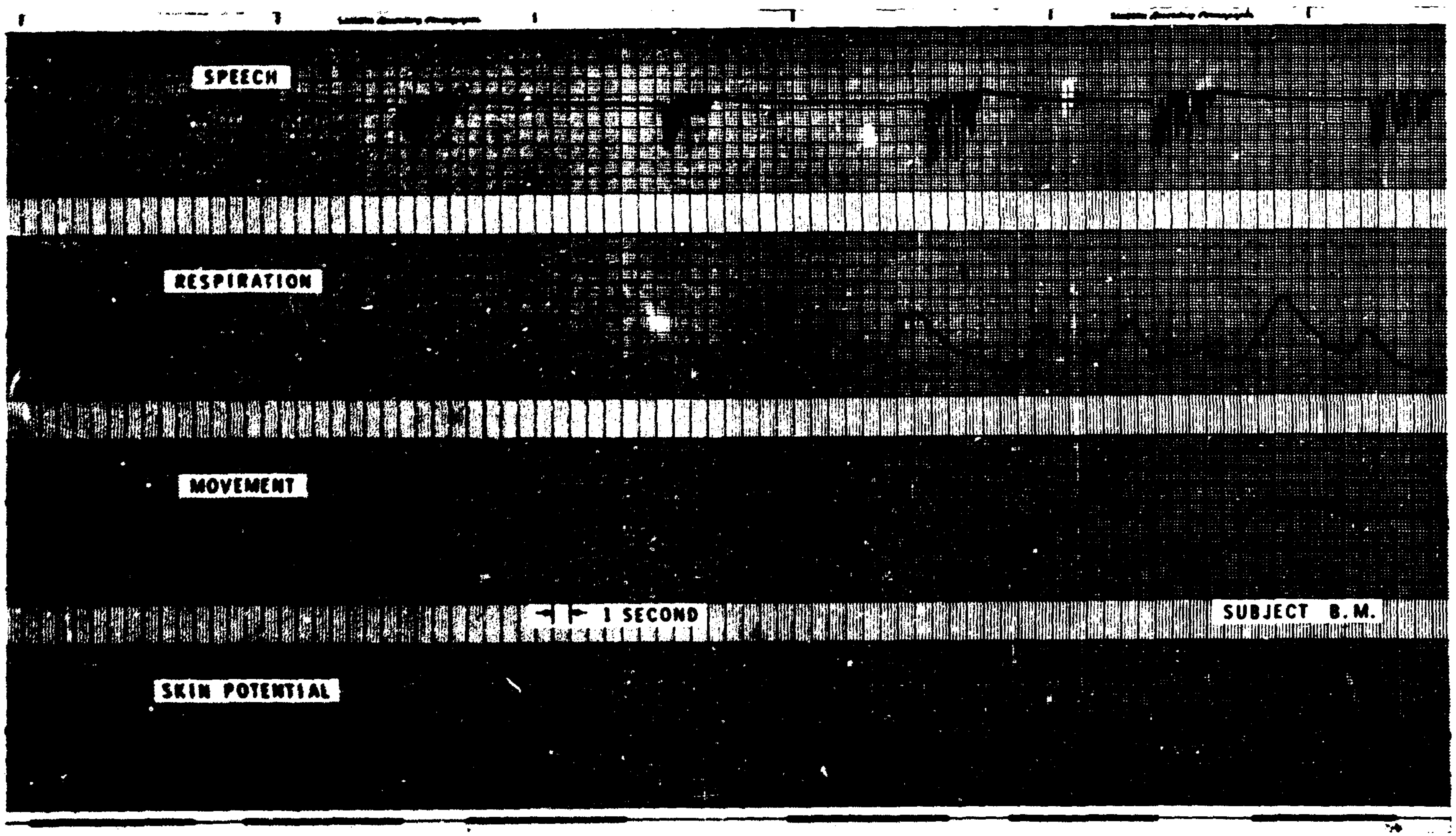


FIG. 5 PORTION OF RECORD OF A SUBJECT'S RESPONSES. DURATION OF EACH METER DISPLAY IS SHOWN IN LOWER MARGIN.

speaks. Movement, as reflected in the third tracing, does not appear to be correlated with the meter displays. Although the tracing of skin potential is seen to change most noticeably during the meter displays, these changes are difficult to interpret and are not representative of all subjects. The response records were intended as a means for uncovering possible physiological changes which could supplement the information gained by direct observation of the subject. They were not intended to provide data for detailed analyses.

A small dynamic microphone (Altec 693A) was used to record the subject's speech. The performance of this microphone was evaluated by comparing its frequency response with the response of a standard condenser microphone (Bruel & Kjaer 4131). Both microphones were placed 7 inches in front of a high-quality 3-inch loudspeaker (KLH Model 6.5) which was mounted in a 6 x 6 x 6-inch wooden enclosure. A sweep oscillator and a level recorder were used to measure the frequency response of the loudspeaker through each microphone. The results of these measurements are shown in Fig. 6.

Since the response of the dynamic microphone was comparable to the response of the condenser microphone in the frequency range 250-7000 Hz, its performance was considered adequate for the recording of speech. Although the fundamental frequency of speech was severely attenuated by the microphone, this important parameter could, of course, still be observed in spectrograms. Prior to each experimental session, the microphone was adjusted to be approximately 4 inches away from, and slightly to the side of, the mouth of the subject. In this position the microphone was out of the way of the subject's breath stream, and the problems frequently encountered with close-talking microphones were thus minimized.

3.3 Experimental Procedure

The experiment consisted of four parts. In the first part, the subject was familiarized with the displays used in the task. He was instructed to read only one meter and to respond by announcing the test word, the meter reading, and again the test word. The displays were not terminated until the subject had responded. The purpose of this part of the experiment was to give the subject experience for the fourth part, which served as the control condition. While the first and fourth parts were identical, the subject was usually much more relaxed during the fourth part.

For the second part of the experiment, the subject was instructed to read all six meters and to announce the sum of his readings. As before, the displays were not terminated until he responded, but the time he required to respond was recorded. This time was averaged over ten displays, and the result was used as the initial display duration in the third and major part of the experiment.

The third part lasted approximately twenty minutes. In this part, the experimenter applied a strategy which was suited to the individual subject. He progressively decreased and occasionally increased the display duration in an effort to establish a stressful situation. All means available were used to make

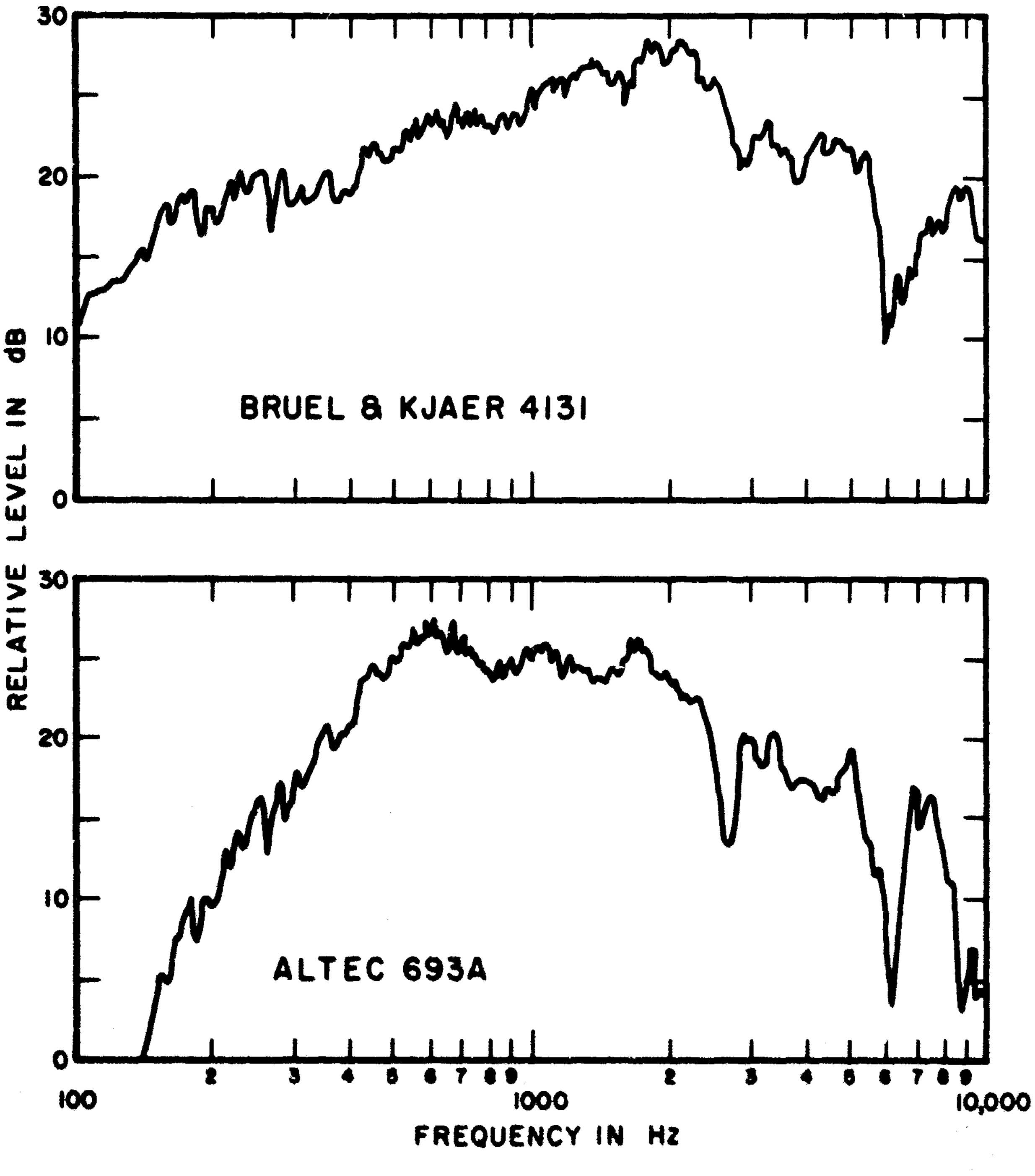


FIG. 6 FREQUENCY RESPONSE CURVES OF CONDENSER MICROPHONE (TOP) AND MICROPHONE USED IN ADDITION TASK (BOTTOM).

inferences about the level of stress induced in the subject. Although the response record from the chart recorder provided useful information, the experimenter tended to rely more heavily on his direct observation of the subject and on listening to the subject's voice. The experimenter was a psychiatrist with extensive experience in recognizing the effects of stress on behavior. For each display, an assistant told the experimenter whether the subject's answer was correct or incorrect. Requests from the subject for an increase in the display duration were also considered to be indicative of stress, but the experimenter did not necessarily comply with such requests.

When the display duration was reduced, most subjects showed signs of frustration and anxiety. Their behavior tended to become disorganized as their performance decreased, and eventually they were unable to read all of the meters before the display was terminated. When they could no longer respond, they appeared to lose interest in the task. At this point, the experimenter increased the display duration until the subject had produced a series of correct responses and thereby regained his confidence. Then the duration was again progressively reduced. These manipulations of the display duration were continued in an effort to stress the subject and maintain his motivation.

Some subjects tended not to respond when they were unsure of their addition. To obtain more samples of their speech under stress, the experimenter might have encouraged these subjects to guess at the answer. But this was not done; the experimenter never spoke to any subject so as not to influence his voice.

There were other subjects for whom it was difficult to create a stressful situation. Reducing the display duration did not bring about any obvious changes in either their behavior or their performance. Whether the difficulties encountered with these subjects were attributable to their personalities, to insufficient motivation, or to shortcomings of the task could not be determined.

Four sets of instructions, corresponding to the four parts of the experiment, were prepared. These instructions are given in the Appendix. Before each part of the experiment, an assistant handed the subject the appropriate set of instructions. In addition, a tape recording of the instructions was played to the subject. The few questions that were asked about the instructions were briefly answered by the assistant. The subjects were paid after the third part of the experiment.

A total of ten subjects participated in these experiments. In the following section of this report, these subjects will be identified only by their initials. All of the subjects were adult males. While most of the subjects were office workers, Subjects D.M., P.M., and J.F. were Air Force pilots on active duty. With the exception of Subject C.W., the subjects were not familiar with the nature or objective of this study. No other criteria were employed in the selection of the subjects.

4. DATA REDUCTION AND RESULTS

The subjects produced an average of 65 veroal responses during the main part of the experiment and about 16 responses during the control condition. Each response consisted of a test word, a numerical answer, and a repetition of the test word. Only a limited number of responses could be analyzed in detail. For each subject, four responses containing a given test word were selected; two of these were obtained while the subject was presumably under stress, and the other two while he was presumably relaxed. Considering the variance normally encountered in repeated utterances of the same speech material, the selection of two responses to represent a given condition was considered mandatory. Nevertheless, the selected data were still minimal, and comparative measurements must be interpreted with caution.

The selected speech samples were analyzed in four ways, three of which produced quantitative results. A paired-comparison listening test was used to determine how reliably the responses obtained under stress could be identified by listeners. Graphic-level tracings were used to measure level changes, and narrow-band spectrograms were used to detect changes in fundamental frequency. The last and most extensive analysis consisted of a qualitative comparison of spectrograms.

4.1 Selection of Responses

The level of stress induced in the subject was not uniform over the entire third part of the experiment. Some responses were therefore more representative of stress than others.

Similarly, in the beginning of the control condition, many subjects were not yet fully relaxed. In order to select sufficiently representative and contrasting responses, it was necessary to make some judgments about the level of stress associated with each response. This was accomplished by listening to the experimenter's recorded comments, by examining the response record obtained with the chart recorder, and also by listening to the subject's speech.

To facilitate the evaluation of these data, a given response record was fastened to the walls of a large listening room. The meter displays indicated on the record for the third and fourth parts of the experiment were numbered consecutively. The two-track tape recording of the experimenter's comments and the subject's speech was presented binaurally to two evaluators who were entirely familiar with these experiments. While listening to the tape, the evaluators followed along from item to item on the response record. In two separate sessions, they rated the level of stress associated with each response on a four-point scale by considering all of the data available to them. The ratings were averaged over the two evaluators and the two sessions, and the results were used to select 20 responses: For each of the two conditions, two responses were selected for each of the five test words.

A few subjects were insufficiently stressed during the third part of the experiment, so that only some of the desired number of stressful responses were available. Other subjects were insufficiently relaxed during the control condition to provide the desired number of control responses. There were two cases in which none of the responses containing a particular test word were considered acceptable representations of one of the conditions: For Subject P.M. no responses containing the test word "Capsule Temperature" were selected, and for Subject C.W. no responses containing the test word "Oil Pressure" were selected.

4.2 Listening Tests

Paired-comparison listening tests were used to determine how well the verbal responses selected to represent the stress and control conditions could be differentiated by listeners. Tape recordings of the selected responses were edited to form ten tests, one test for each subject. Most tests consisted of ten pairs of verbal responses; the responses of a given pair featured the same test word but represented different conditions. Each of the five test words appeared equally often. The order in which the test words were sampled, and the position of the stressful response in each pair, followed a random pattern. This pattern varied from test to test. For those subjects who produced only some of the desired responses, the tests contained fewer than ten pairs of responses.

The tests were presented at a comfortable listening level to a group of nine listeners over a high-quality loudspeaker system. Each test was administered three times in succession, and the listeners were provided with answer sheets which were individually collected after each presentation. The listeners were instructed to mark each test item with the number 1 or with the number 2, depending on whether they thought that the first or the second speech sample of the pair was more indicative of stress. Seven of the listeners were also subjects, so that they heard their own speech when their tests were presented.

The number of correct ratings obtained from a given listener for a given test was divided by the total number of ratings for that test (usually 30) to derive a listener score. This score indicated how well the listener could identify the stressful responses of a particular subject. Similarly, the number of correct ratings obtained from all listeners for a given item (response pair) of a given test was divided by 27, the total number of ratings from all listeners for that item, to derive a test-item score. This score indicated how well the group of listeners could identify the stressful response in a particular test item. The listener scores and the test-item scores, expressed in percent, are given in Tables I and II, respectively.

As indicated in Table I, the listener scores were averaged separately over the ten subjects and over the nine listeners. The listener means, which range from 66 to 83 percent, suggest that the individual listeners were about equally capable of identifying the stressful responses of all subjects. The subject means, on the other hand, have a much greater range and clearly show that the listeners were more successful with some subjects than with others. For example, while they could readily identify the stressful responses of Subjects J.C., C.W., and J.F., their performance level was close to chance (50 percent) for Subjects R.E., D.M., and P.M. The degree to which individual subjects verbally communicated stress to the listeners varied considerably.

The rectangle within Table I encloses the scores obtained from those listeners who were also subjects. The scores obtained from listeners who heard their own speech, which are shown along the diagonal, generally exceeded the corresponding subject means.

TABLE I

Listener scores (in percent) indicating how well each listener could identify the stressful responses of each subject.

Listener	Subject										Mean
	J.C.	R.B.	B.M.	R.E.	D.K.	B.E.	C.W.	D.M.	P.M.	J.F.	
J.C.	90	70	93	47	81	67	92	60	33	93	73
R.B.	80	90	67	40	57	80	79	53	59	83	69
B.M.	100	83	93	53	57	87	100	83	37	100	79
R.E.	97	77	74	87	62	80	92	60	63	97	79
D.K.	93	97	82	100	90	83	92	70	25	100	83
B.E.	93	83	74	80	57	63	100	70	54	87	76
C.W.	100	97	56	50	71	60	100	77	25	100	74
M.H.	100	63	74	77	76	87	96	67	79	90	81
K.S.	100	60	74	43	62	70	96	37	17	100	66
Mean	94	80	77	64	68	75	94	64	44	94	

TABLE II

Test-item scores (in percent) indicating how well all listeners could identify the stressful response in each test item.

	Subject									
Test Word	J.C.	R.B.	B.M.	R.E.	D.K.	B.E.	C.W.	D.M.	P.M.	J.F.
Generator	100	33	63	48	93	44	100	70	56	100
Current	82	89	93	70		70	100	52	33	96
Capsule	93	82	70	30	59	89	78	15		82
Temperature	96	89		63		67	85	67		96
Oil	93	100	44	70	78	78		85	74	96
Pressure	100	100	89	63	56	96		70	30	96
Deviation	96	74	78	93	52	100	100	85	93	100
Angle	100	96	85	70	59	89	100	74	19	78
Relative	89	56	70	70	93	52	96	44	30	100
Velocity	96	85	93	78		67	96	82	17	96

However, the responses of a given subject were rarely better differentiated by that subject than by any other listener. Also, a given listener was not necessarily a better judge of his own speech than he was of the speech of other subjects.

The test-item scores shown in Table II are arranged according to the test words which appeared in the response pairs. For most subjects, two scores are shown for each test word because a given test word usually appeared in two response pairs. The listeners, of course, rated each pair by considering the responses in their entirety and not just those portions containing the test words. If scores falling outside of the range 35-65 percent are arbitrarily regarded as being significantly different from chance, more than two-thirds of the scores shown indicate that the response pairs reflected the stressed and relaxed states of the subjects in the expected manner. Only eight scores are lower than 35 percent. These low scores mean that the responses obtained during the control condition were rated to be more stressful than the responses obtained during the third part of the experiment.

Some of the low scores can be explained in terms of the original selection of the control responses. The evaluators who estimated the level of stress induced in a given subject at the time each response was uttered were occasionally confronted with conflicting information; the experimenter's comments may have implied that the subject was already relaxed, but his speech may have still sounded stressful. Thus, a few stressful responses could have been accidentally selected as control responses. Other responses obtained during the control condition may have contained gross features which led the listeners to believe that these responses were produced under stress. An example of such features is a correction in the numerical answer: "Oil Pressure, plus -- minus forty-five, Oil Pressure."

4.3 Level Measurements

Because the recording microphone was attached to the subject's head and was thus kept at a constant distance from his mouth, it was possible to measure changes in the level of his speech. Such level changes were regarded as potential indicators of stress. In view of the small number of responses obtained for each condition, it was considered necessary to restrict the measurements to corresponding portions of responses containing the same test word. Graphic-level tracings were prepared from all of the selected responses by a given subject. These tracings were then used to read the levels of certain portions of the test words occurring in the initial and final positions of the responses. The central portions of the responses, which contained the subject's numerical answers, were disregarded.

Typical graphic-level tracings of the five basic responses are shown in Fig. 7. Of the various peaks traced out for each of the five test words, only two were selected for level measurements. These peaks, which were associated with specific portions of each test word (two phonetically stressed vowel nuclei), could be readily identified in the tracings for all subjects. Lines were drawn through the selected peaks at integral decibel values. For a given test word, the four level measurements were

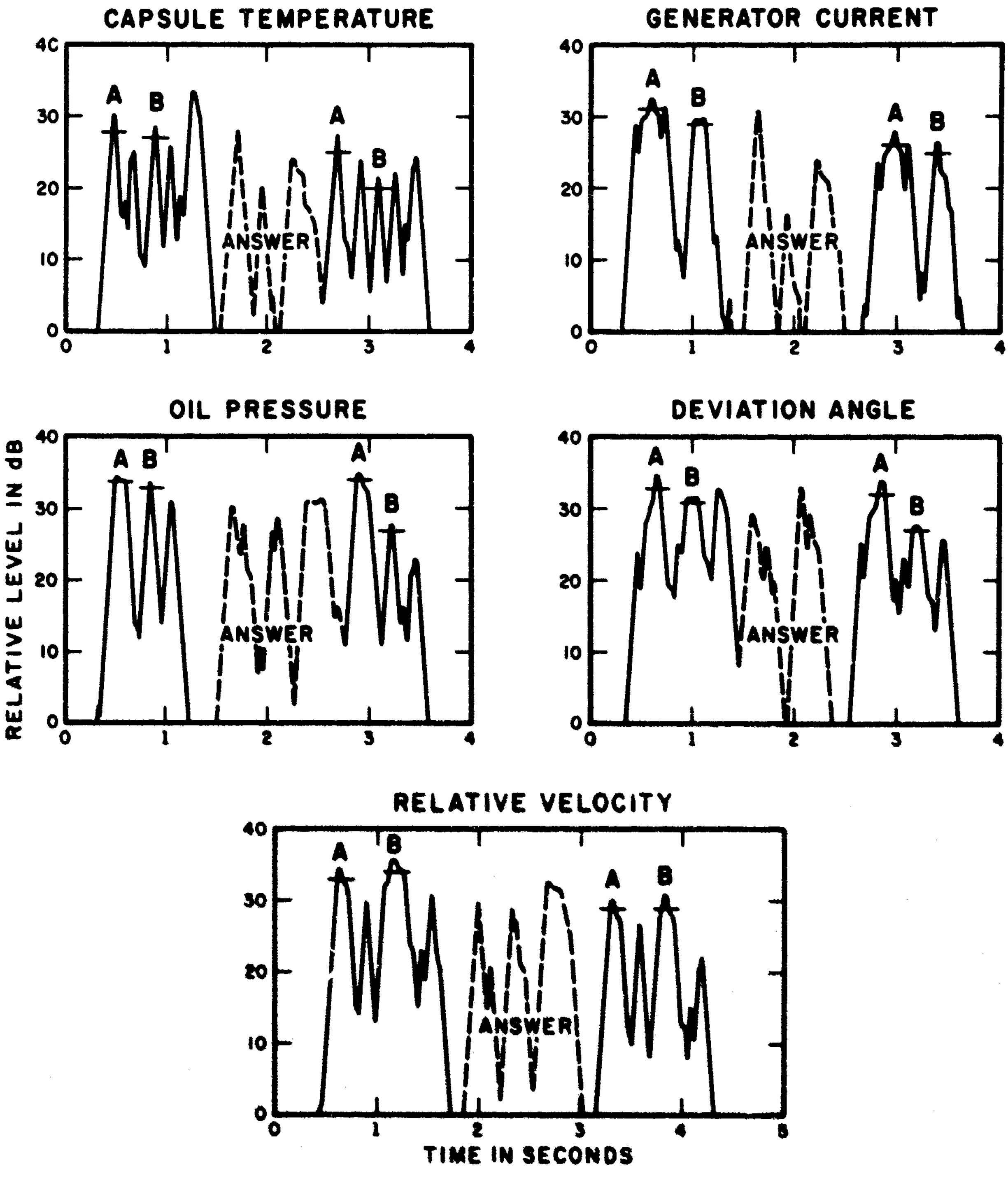


FIG. 7 TYPICAL GRAPHIC-LEVEL TRACINGS OF THE FIVE BASIC RESPONSES. TWO PORTIONS OF EACH TEST WORD (LABELED A AND B) WERE SELECTED FOR LEVEL MEASUREMENTS.

individually averaged over the two responses representing the same condition. The means representing the control condition were then subtracted from the means representing the stress condition. The resulting data are given in Table III.

This table shows that some subjects produced relatively large differences which were fairly consistent for all test words. Subject R.E. always spoke more softly when he was under stress. When Subjects B.M. and P.M. were stressed, they always spoke the initial portion of their responses more softly; their level change during the final portion depended somewhat on the particular test word. Subject D.K. generally spoke louder under stress. The level changes recorded for the remaining subjects show only minor trends or inconsistencies from measurement to measurement. No data were available for Subject C.W.

4.4 Measurements of Fundamental Frequency

Changes in the fundamental frequency of a person's voice are often thought of as manifestations of stress. From a physiological point of view, a person under stress may generate a higher pressure in his lungs (possibly to provide greater ventilation), and this could give rise to a higher fundamental frequency. He may also have less control over the muscles that operate his larynx. The tension of his vocal folds may vary, and this could change their frequency of vibration. To determine whether the subject's responses obtained in the third part of the experiment and in the control condition differed with respect to fundamental frequency, narrow-band spectrograms were prepared and used to measure the average fundamental frequency of all voiced portions of the test words.

TABLE III

Differences (in dB) between mean levels obtained for selected portions of test words spoken in control condition and under stress. A negative difference indicates that stress produced the lower mean level.

Test Word	Word Pos.	Word Port.	Subject J.C.	R.B.	B.M.	R.E.	D.M.	P.M.	J.F.	D.K.	B.E.
Generator Current	I	A	2.0	0	-2.5	-4.5	-2.0	-4.5	-1.5	3.0	0
	I	B	2.0	0.5	-0.5	-4.5	-1.5	-2.0	1.5	2.0	-1.0
	F	A	5.0	1.5	0	-4.0	-1.0	-6.5	3.0	2.0	-4.0
	F	B	1.0	0.5	1.0	-4.5	-0.5	-5.0	4.0	-1.0	-5.5
Capsule Temperature	I	A	-4.5	-4.0	-6.0	-5.5	-7.5		-4.5	2.0	-2.0
	I	B	2.5	0.5	-1.0	-7.0	0.5		5.0	4.0	-0.5
	F	A	2.0	0.5	-3.0	-5.5	2.5		-1.5	3.0	-1.5
	F	B	3.5	-1.5	-3.0	-5.5	2.0		0	5.0	7.5
Oil Pressure	I	A	-0.5	0.5	-1.0	-2.0	-0.5	-1.5	1.0	2.5	-4.0
	I	B	1.0	0	-3.5	-2.5	0.5	-1.5	3.0	1.0	1.5
	F	A	0.5	0	0	-2.5	1.5	3.5	2.5	1.0	-3.0
	F	B	-4.0	4.0	3.0	-1.5	-1.5	3.5	-2.0	0.5	-3.5
Deviation Angle	I	A	0	0	-3.0	-5.0	-1.5	-3.0	0	-1.0	0
	I	B	2.5	-1.5	-2.5	-5.5	2.5	-2.5	0.5	1.0	1.5
	F	A	3.5	2.0	-2.0	-4.5	1.5	-3.5	0.5	-0.5	-2.5
	F	B	3.5	-2.5	-1.0	-3.5	0	-3.0	2.0	0	-1.5
Relative Velocity	I	A	-7.0	2.0	-3.5	-4.0	-6.5	-1.5	2.0	5.0	-1.0
	I	B	0.5	1.5	-2.0	-5.0	-2.0	-2.0	-2.0	3.0	0
	F	A	3.0	0	-2.0	-2.0	-1.0	-3.5	0	3.0	-3.0
	F	B	1.5	1.5	-2.0	-6.0	-2.5	0	-5.0	4.0	-2.0
All	I		-0.2	-0.1	-2.6	-4.6	-1.8	-2.3	0.5	2.3	-0.6
	F		2.0	0.6	-0.9	-4.0	0.1	-1.8	0.4	1.7	-1.9

Since the duration of a typical response exceeded the time interval encompassed by the spectrograph (2.4 seconds), separate spectrograms were prepared for the test words occurring in the initial and final positions of the response. The measurements of fundamental frequency were taken by reading the frequency of the fifth harmonic and dividing the result by five. The average fundamental frequency of each vocalic interval throughout each test word was thus determined. There were some cases in which the vibration of the vocal folds was too irregular or too low in frequency to provide a readable harmonic structure in the spectrogram. In other spectrograms, the harmonic structure was obscured by noise from excessive aspiration. No measurements of fundamental frequency could be obtained for these cases.

For a given test word in a given position (initial or final), the measurements obtained from the two spectrograms that represented the same condition were individually averaged. The means representing the control condition were then subtracted from the means representing the stress condition. The results of these calculations are presented in Table IV. For each test word, several difference values are listed; these correspond to the sampled vocalic intervals.

A number of trends are evident in the data of Table IV. Subject J.C. always produced a higher fundamental frequency when he was under stress, while Subject R.E. always produced a lower fundamental frequency. Subject P.M. tended to lower his fundamental frequency. It is of interest that Subjects R.E. and P.M. also tended to speak more softly when they were stressed. It might be supposed, therefore, that these subjects used a lower

TABLE IV

Differences (in Hz) between mean values of fundamental frequency obtained for various portions of test words spoken in control condition and under stress. A negative difference indicates that stress produced the lower mean value.

Test Word and Position →	Subject J.C. I	J.C. F	R.B. I	R.B. F	B.M. I	B.M. F	R.E. I	R.E. F	D.M. I	D.M. F
Generator Current	13.6 19.0	12.2	1.4 3.4	5.4 0.0	-4.7 5.4 1.4	2.0	-4.7 -5.4	-1.4 -14.9	4.7 8.1	-4.7
Capsule Tempera-ture	2.0 10.2 18.3 12.2	13.6 14.9 6.8	-8.8 2.0 0.0 -5.4	0.0 1.4 3.4	0.0 6.8 13.6 20.4 13.6	0.0 -6.8 -3.4 -17.0	-13.6 -15.6 -19.0 -3.4	-14.9	13.6 4.7 17.0 17.0	8.8 6.8
Oil Pressure	8.1 11.5 14.9	15.6	6.8 12.2 20.4	3.4 3.4	-15.6 6.8 14.9	2.0	-3.4 0.0	-18.3	-2.0 19.0	15.6
Deviation Angle	13.6 19.0	14.9 2.0	5.4 -8.1	-3.4 3.4 -6.8	-4.7 5.4	6.8 6.8	-2.0 -3.4		-4.7 39.0	14.9 0.0
Relative Velocity	4.7 13.6 18.3	6.8 1.4	10.2 3.4 3.4	3.4 5.4	4.7 10.2 3.4	0.0 -4.7 1.4	-6.8 -8.1	-6.8 -3.4	-3.4 6.8 -6.8 5.4	2.0 -13.6
All	12.8	9.8	3.3	1.6	5.1	-1.2	-7.1	-10.0	8.5	3.7

TABLE IV (cont.)

Test Word and Position →	Subject									
	P.M.		J.F.		D.K.		B.E.		C.W.	
	I	F	I	F	I	F	I	F	I	F
Generator Current	-5.4 -8.1	-4.7 -10.2	-11.5 -6.8		13.6 23.7 3.4	6.8 -6.8 -3.4	-17.0 4.7	-10.2 -4.7	1.4 -8.1	0.0 2.0
Capsule Temperature			-13.6 -10.2		0.0 6.8 3.4 3.4	6.8 -6.8 -10.2 -3.4	-10.2 -10.2 6.8 13.6 17.0	5.4 10.2 1.4		3.4 -10.2 -1.4 -6.8
Oil Pressure	6.8 18.3 6.8	10.2	14.9	-2.0	8.1 12.2 -1.4	-1.4 -3.4	-23.7 19.0 34.0	-17.0 1.4		
Deviation Angle	-12.2 -10.2	-1.4	8.8 -17.0	-4.7 -4.7	3.4 4.7 10.2	-4.7 -1.4 0.0	-37.3 15.6	-2.0 8.8	2.0 4.7	-10.2 -3.4
Relative Velocity	-3.4 -12.2 -5.4 -23.7	-4.7 -10.2 -5.4	10.2 -3.4 -2.0		10.2 3.4 3.4 0.0	6.8 6.8 3.4	-21.7 -15.6 4.7	-2.0 -8.8		-5.4 -2.0 3.4 3.4
All	-4.4	-3.8	-3.1	-3.8	6.4	-0.7	-1.3	-1.6	0.0	-2.3

lung pressure when they generated speech under stress. For Subject J.F., only three measurements could be obtained for test words in the final position; most of his narrow-band spectrograms of test words in this position were too irregular to be read. However, his corresponding wide-band spectrograms showed that his fundamental frequency was generally lower when he was stressed. Subject D.K. increased his fundamental frequency for test words in the initial position. At the beginning of a test word in the initial position, Subjects B.M. and B.E. tended to produce a lower fundamental frequency, but toward the end of the word they always produced a higher fundamental frequency. For these subjects, there were differences in the contour of the fundamental frequency between the control condition and the stress condition. These differences are discussed further and illustrated in the next section.

4.5 Comparison of Spectrograms

Level and fundamental frequency are only two parameters of speech which may be affected by stress. However, the effects of stress on other speech parameters are generally more difficult to define and measure. A person under stress may devoice and aspirate certain portions of his speech. Also, he may be less precise in his articulation; he may inadvertently slur syllables together and omit some speech sounds altogether. While such effects are difficult to quantify, they can be readily demonstrated with the aid of spectrograms.

Wide-band and narrow-band spectrograms were prepared of all test words contained in the selected responses. Separate spectrograms were made of the test words occurring in the initial and final positions of each response. For a given subject, test word, and test-word position, the two spectrograms representing the control condition were paired with the two spectrograms representing the stress condition. The paired spectrograms were critically compared. Only those differences which were evident in both pairs of spectrograms were attributed to stress; differences found in only one pair were disregarded. Various effects of stress were observed in these comparisons. The effects most consistently demonstrated by each of the ten subjects are illustrated in Figs. 8 through 16.

These figures show mostly wide-band spectrograms made with a linear 7-KHz frequency scale. Some figures also show narrow-band spectrograms made with a linear 3.5 KHz frequency scale, contour spectrograms, and oscilloscope photographs. The spectrograms are always arranged in pairs; the spectrograms representing the control condition are shown on the left side, and the corresponding spectrograms representing the stress condition are shown on the right. The subjects and the test words are identified in the figures. The symbols (I) and (F) indicate respectively whether the test word was in the initial or final position of the response. In almost all cases, the subjects uttered the complete response (i.e., test word, numerical answer, test word) on one expiratory cycle. Thus, spectrograms of test words in the final position usually show a falling fundamental frequency at the end of the utterance.

The manifestations of stress are seen to vary considerably from subject to subject. It is appropriate, therefore, to discuss the data for each subject separately. There are, however, some characteristics that appear for several subjects, and these characteristics will be identified and described first.

Changes in fundamental frequency are best observed and measured from narrow-band spectrograms, as was noted in Section 4.4. However, wide-band spectrograms also provide an indication of the fundamental frequency. Since a wide-band spectrogram shows the individual glottal pulses as vertical lines, the distance between successive lines is, of course, a measure of the fundamental period. For portions of utterances in which the fundamental frequency is very low, or in which there are considerable fluctuations in the interval between glottal pulses, wide-band spectrograms provide the best indication of laryngeal activity. Thus, for example, the upper right spectrogram in Fig. 8 shows a very long interval between pulses at the end of the utterance. An extended and irregular glottal period can also be seen in the right-hand spectrograms in Fig. 13.

A characteristic that appears, with slight variation, in the stressful utterances of several subjects is an irregularity in the pattern of glottal vibration. This effect, which will be referred to as voicing irregularity, is exemplified in the right-hand spectrogram of the test word "Deviation Angle," shown in Fig. 16. The voicing irregularity is most marked in the time intervals 1.5 and 1.7 sec, and 1.8 to 2.0 sec. A considerable variation in the darkness of successive vertical lines is observed, particularly at high frequencies. The spectral peaks associated with the higher-order formants fluctuate in level from one glottal

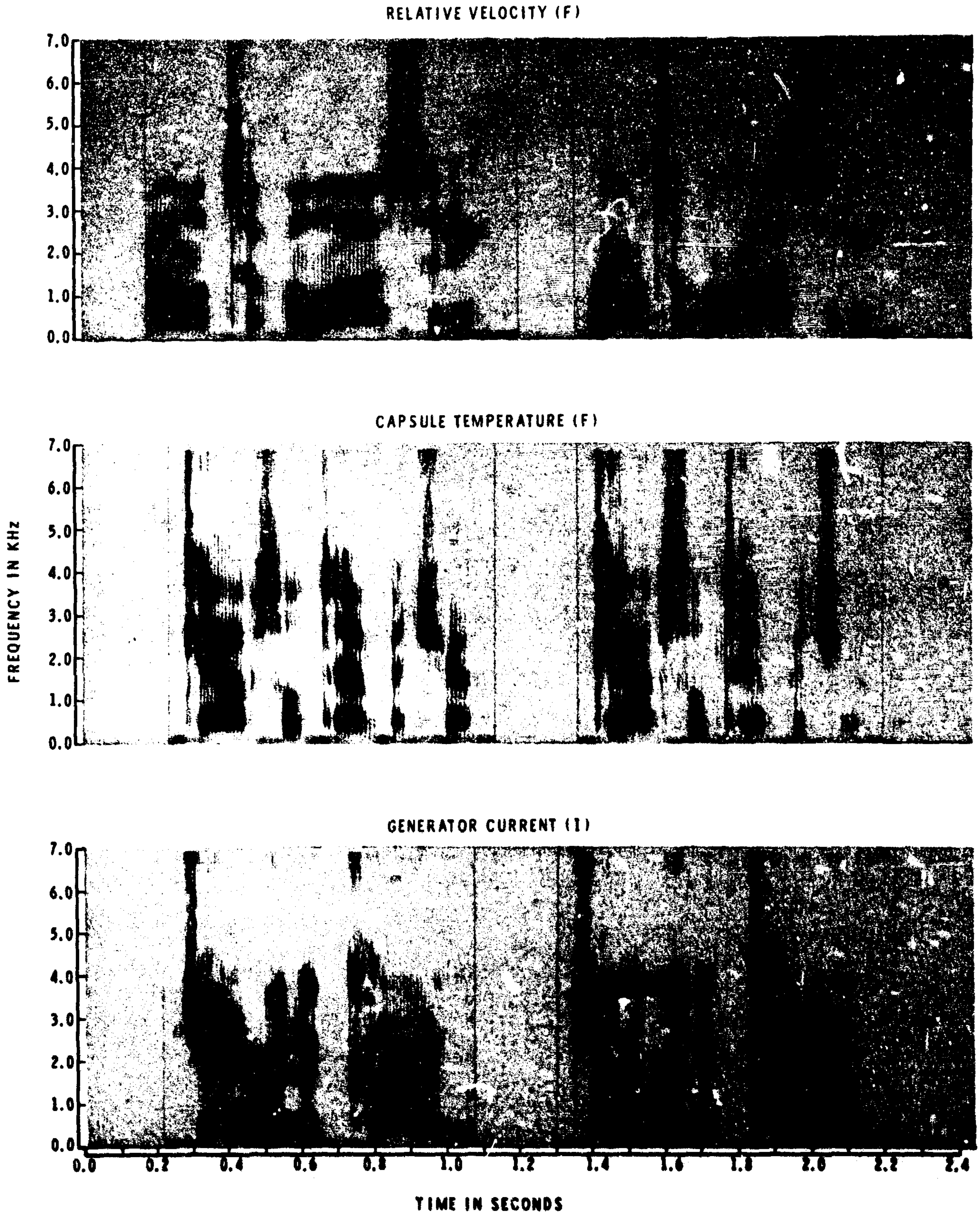


FIG. 8 SPECTROGRAMS OF TEST WORDS SPOKEN BY SUBJECT J.C. IN CONTROL CONDITION (LEFT) AND UNDER STRESS (RIGHT).

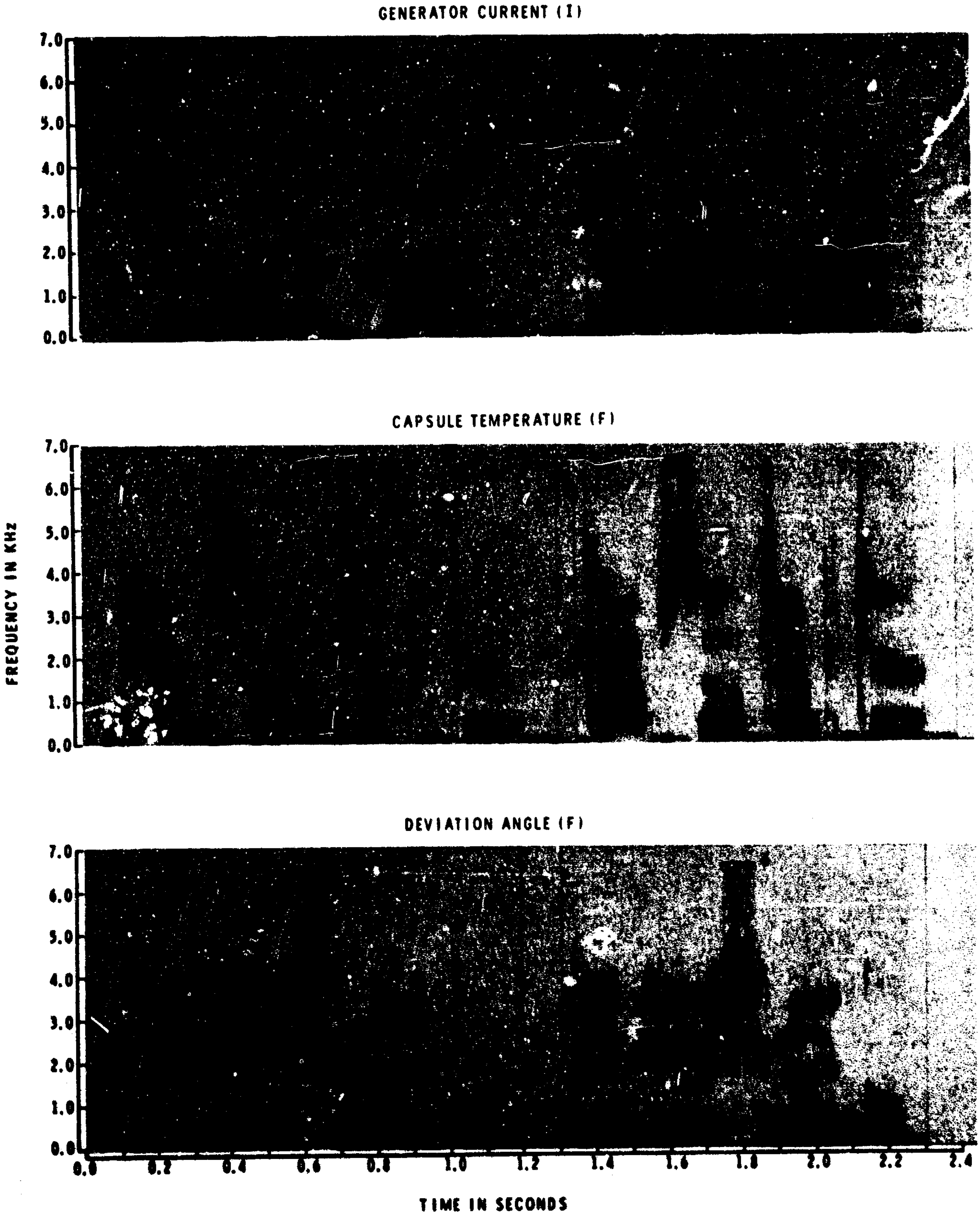


FIG. 9 SPECTROGRAMS OF TEST WORDS SPOKEN BY SUBJECT R.B. IN CONTROL CONDITION (LEFT) AND UNDER STRESS (RIGHT).

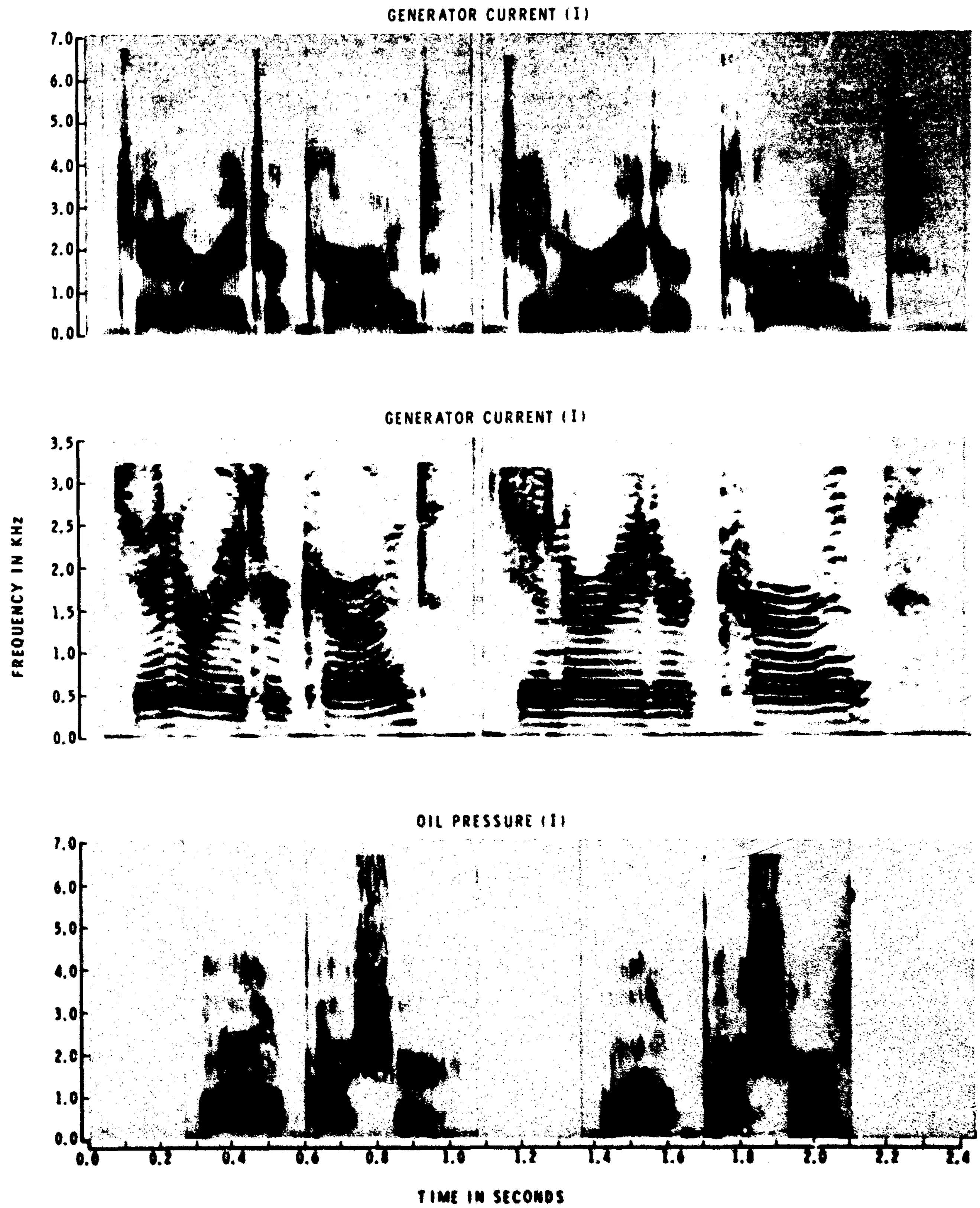


FIG. 10 SPECTROGRAMS OF TEST WORDS SPOKEN BY SUBJECT B.M. IN CONTROL CONDITION (LEFT) AND UNDER STRESS (RIGHT).

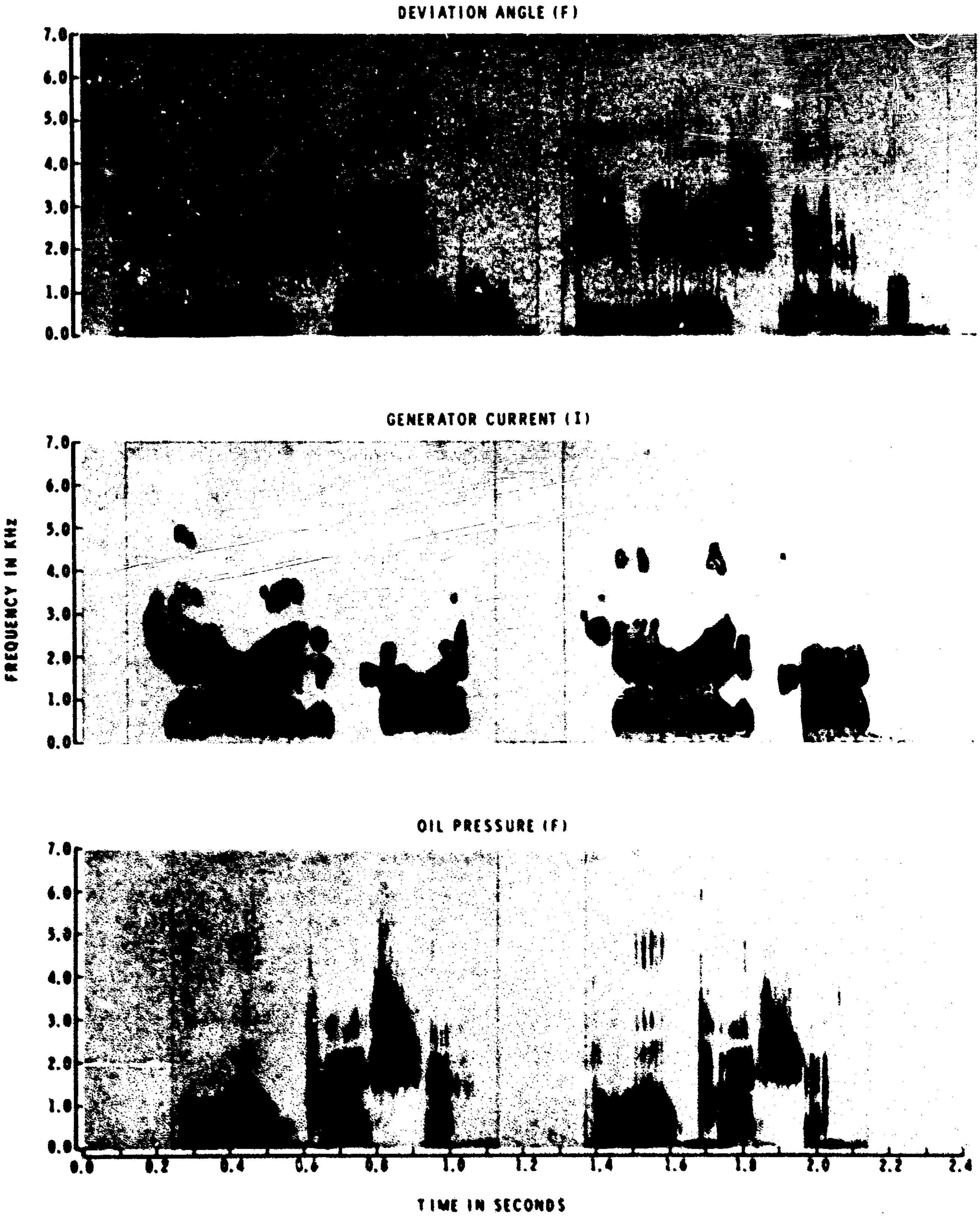


FIG. 11 SPECTROGRAMS OF TEST WORDS SPOKEN BY SUBJECT R.E. IN CONTROL CONDITION (LEFT) AND UNDER STRESS (RIGHT).

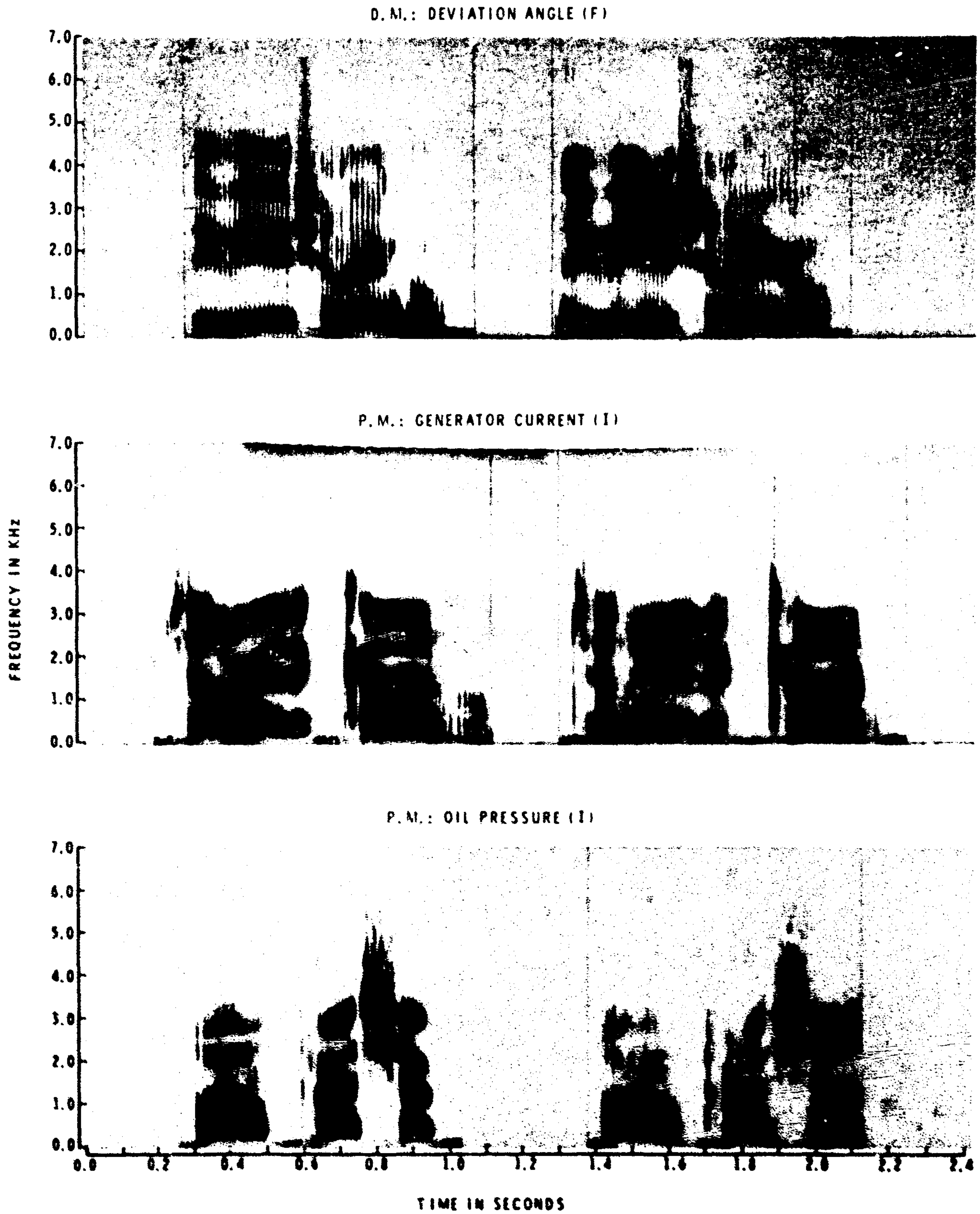


FIG. 12 SPECTROGRAMS OF TEST WORDS SPOKEN BY SUBJECTS D.M. AND P.M. IN CONTROL CONDITION (LEFT) AND UNDER STRESS (RIGHT).

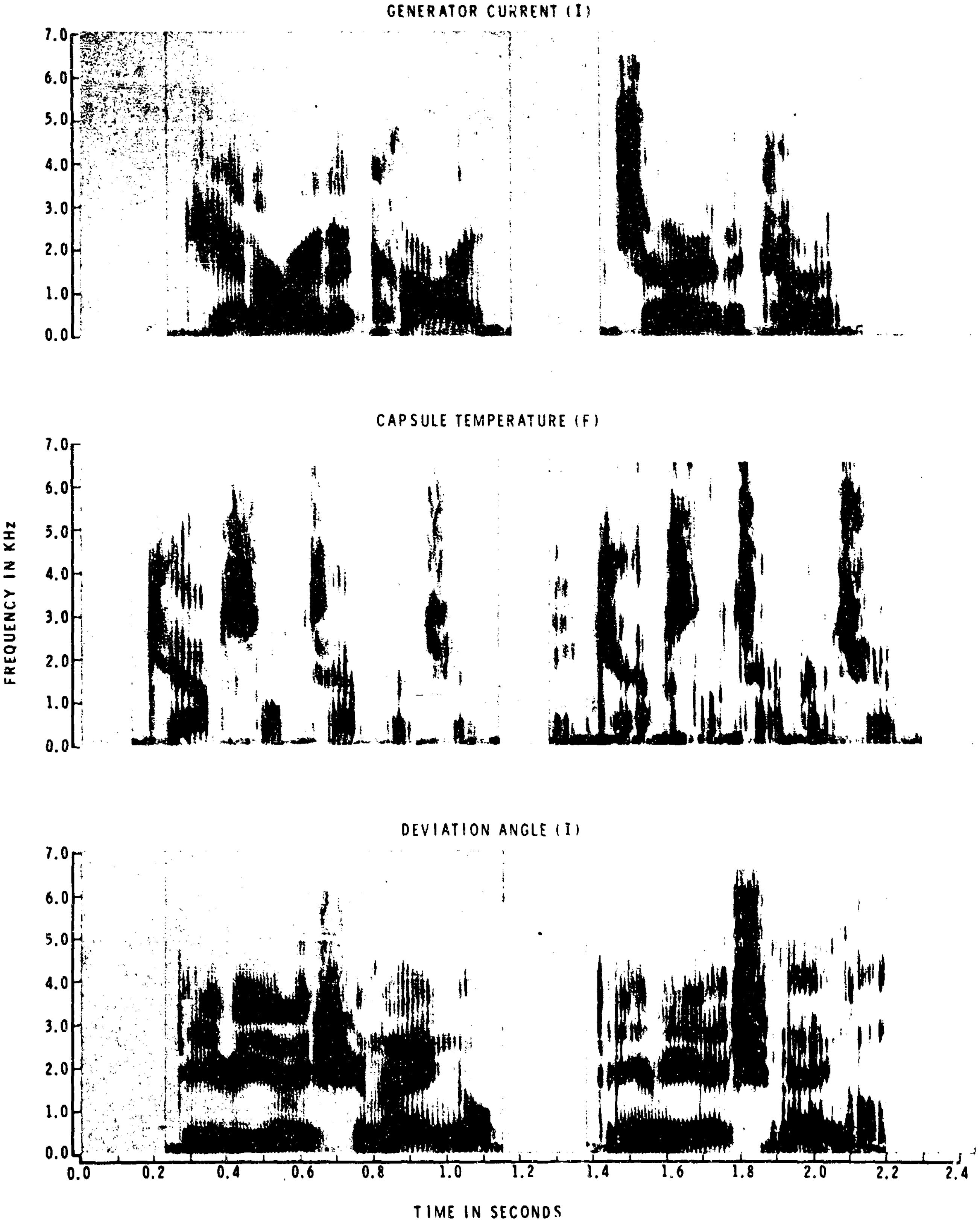


FIG.13 SPECTROGRAMS OF TEST WORDS SPOKEN BY SUBJECT J.F. IN CONTROL CONDITION (LEFT) AND UNDER STRESS (RIGHT).

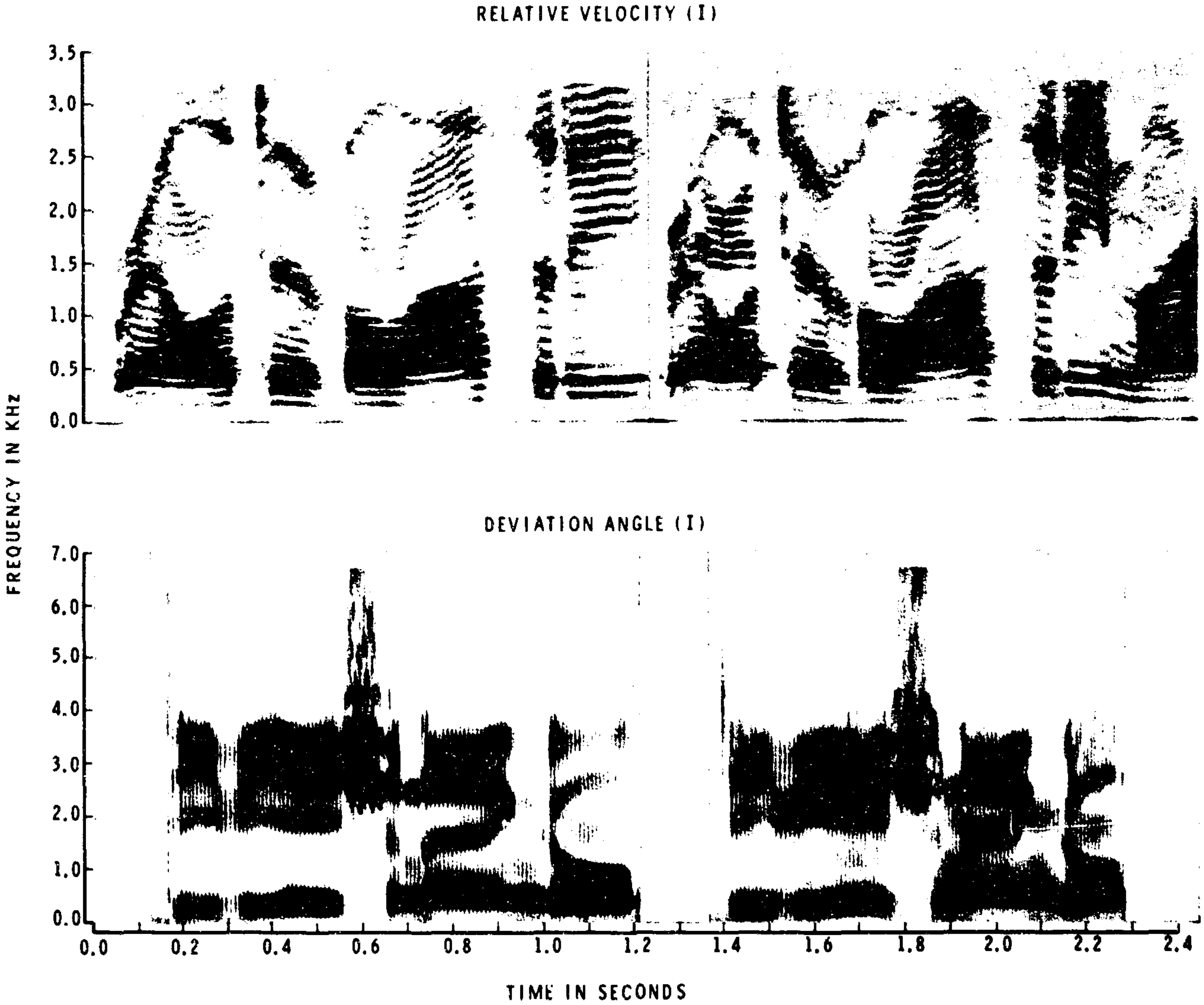


FIG.14 SPECTROGRAMS OF TEST WORDS SPOKEN BY SUBJECT D.K. IN CONTROL CONDITION (LEFT) AND UNDER STRESS (RIGHT).

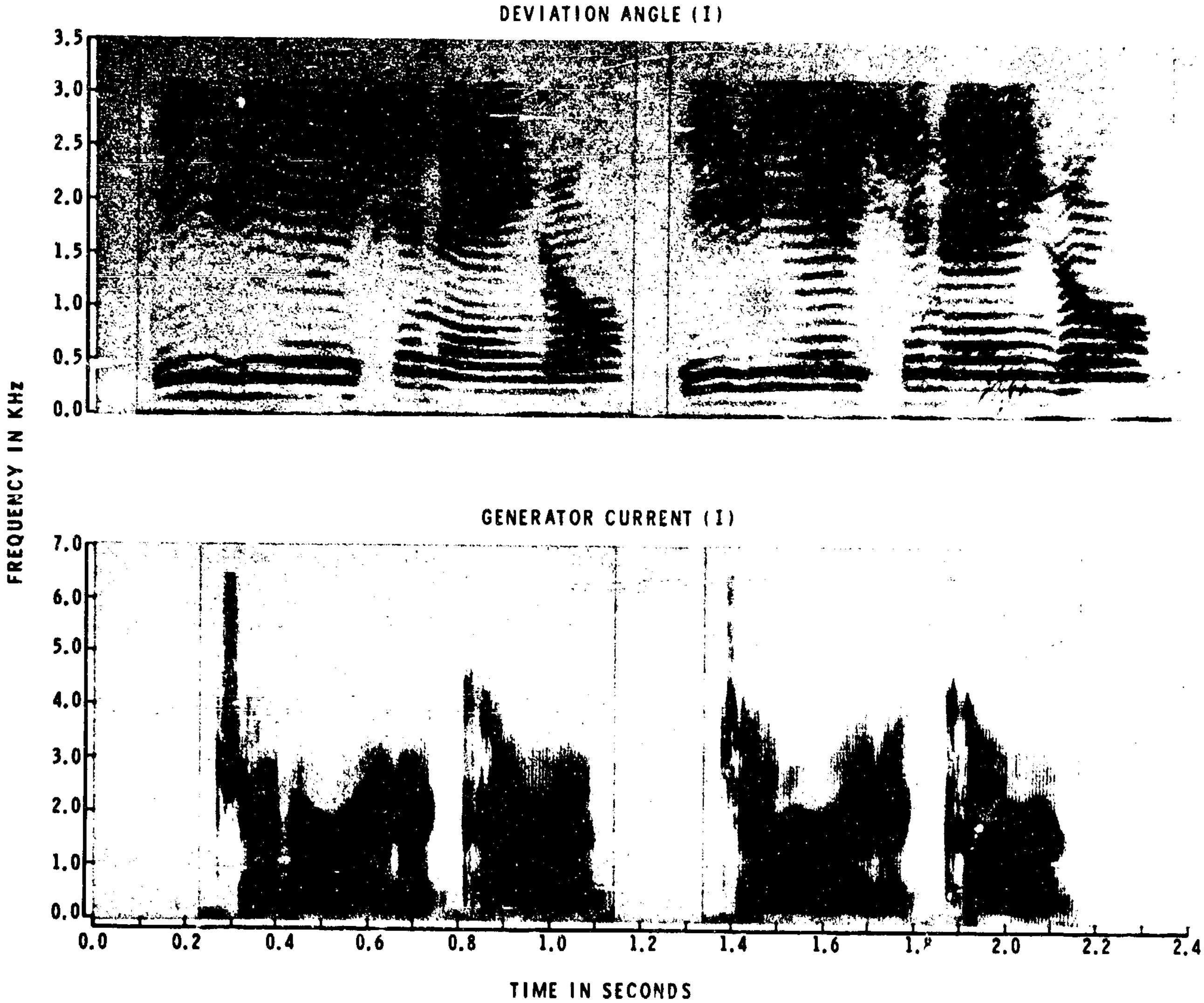


FIG. 15 SPECTROGRAMS OF TEST WORDS SPOKEN BY SUBJECT B. E. IN CONTROL CONDITION (LEFT) AND UNDER STRESS (RIGHT).

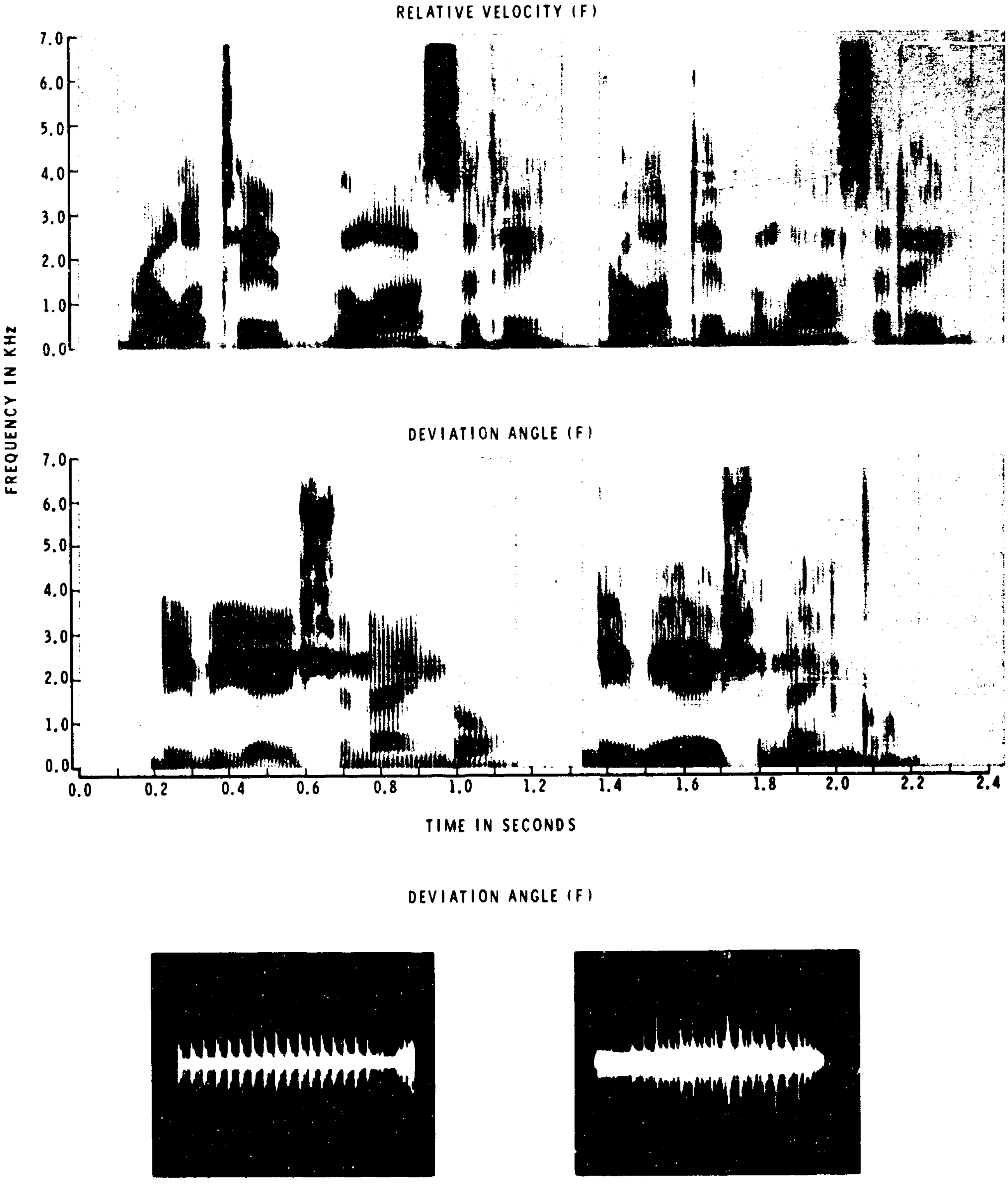


FIG. 16 SPECTROGRAMS AND OSCILLOSCOPE PHOTOGRAPHS OF TEST WORDS SPOKEN BY SUBJECT C.W. IN CONTROL CONDITION (LEFT) AND UNDER STRESS (RIGHT).

pulse to the next. In the corresponding spectrogram representing the control condition, the vertical lines have a uniform appearance. Voicing irregularity presumably occurs when successive glottal pulses have different shapes and, hence, different spectra. Minor differences in shape would have a greater influence on the spectrum at high frequencies than at low frequencies. Such perturbations could arise as a consequence of excessive mucus on the vocal folds, or possibly as a consequence of inadequate control of the laryngeal musculature.

This effect can be more clearly demonstrated with the aid of the oscilloscope photographs shown in Fig. 16. These waveforms correspond to the spectrograms of the test word "Deviation Angle" and cover the time intervals 0.4 to 0.6 sec and 1.5 to 1.7 sec. The speech signal was processed by a 1.5 - 2.5 KHz bandpass filter. Successive glottal pulses show considerable differences in amplitude in the right-hand waveform, whereas there is greater uniformity in the left-hand waveform. Other examples of voicing irregularity appear in Fig. 10 (upper right spectrogram, [ɛ] in "Generator"), in Fig. 13 (lower right spectrogram), and elsewhere.

Another characteristic is the relative amount of high-frequency energy in the glottal pulses. For some subjects, the spectrograms representing the stress condition show less energy in the frequency region of the third and fourth formants than the spectrograms representing the control condition. This characteristic is usually more evident in front vowels than in the back vowels because front vowels tend to have more high-frequency energy. In the spectrograms of the test word "Relative

Velocity" shown in Fig. 8, for example, the third and fourth formants in the first two syllables of "velocity" are weaker for the stress condition than for the control condition. Similar differences are observable in Fig. 12 (spectrograms at bottom) and in Fig. 13 (spectrograms at bottom).

Stress also appears to influence the manner in which a talker controls the movements of his articulatory structures. Some subjects, when they are stressed, have a tendency to generate certain consonants and vowels more rapidly, often with a less constricted vocal tract. In a sense, their articulatory movements for a particular phonetic segment fall short of the intended target. This effect can often be observed in the test word "Generator Current"; the consonants [n] and [t] in "Generator" seem to be particularly susceptible to distorted or less precise articulation. Examples are seen in Fig. 9 (upper spectrograms), Fig. 13 (upper spectrograms), and Fig. 15 (lower spectrograms). Examples of less precise articulation of vowels are seen in the spectrograms of the test word "Oil Pressure" shown in Fig. 11, and in the spectrograms of the test word "Generator Current" shown in Fig. 13.

The data for the individual subjects will now be discussed in detail. Except when noted otherwise, all comments will be made with respect to the right-hand spectrograms, which represent the stress condition.

Subject J.C.

The effects of stress on the speech of this subject are illustrated in Fig. 8. In the spectrogram of "Relative Velocity," there is a marked drop in fundamental frequency in the last voiced interval following the [s]. Voicing irregularity also appears in this interval. The spectrogram also exhibits less high-frequency energy in the glottal pulses. The [s], on the other hand, contains more high-frequency energy. Extreme devoicing of the last syllable is evident in the spectrogram of "Capsule Temperature." In the spectrogram of "Generator Current," the subject did not produce a smooth transition from the initial vowel into the nasal consonant. This suggests that when the subject is under stress he has less control of the position of the velum.

Subject R.B.

The spectrograms for this subject are shown in Fig. 9. This subject was less precise in his articulatory movements when he was stressed; he tended to slur over, or even omit, certain consonants. In the spectrogram of "Generator Current," the initial [n] is almost missing, the duration of the stop closure for [t] in "Generator" is reduced almost to zero, and there is no tight closure for [k] in "Current." In the spectrogram of "Capsule Temperature," several of the vowel durations and stop gaps are shortened. The left-hand spectrogram of "Deviation Angle" shows that this speaker tended to exhibit more voicing irregularity in the control condition. This result demonstrates that voicing irregularity does not always occur as a consequence of stress.

Subject B.M.

This subject spoke in a rather precise fashion in the control condition, with careful articulation of all syllables. Under stress, however, he had a tendency to alter the timing in portions of his utterances. The wide-band spectrogram of "Generator Current," shown in Fig. 10, indicates that he slurred the [t] in "Generator" and aspirated the voiceless [k] and [t] in "Current" more heavily. There also is some voicing irregularity in the initial [ɛ] in "Generator." In the control condition, this subject spoke with considerable inflection, as suggested by the left-hand narrow-band spectrogram. The right-hand narrow-band spectrogram shows less pitch fluctuation; the subject tended to speak in more of a monotone when under stress. The left-hand spectrogram of "Oil Pressure" shows some voicing irregularity in the control condition. In the listening tests, the response containing this particular utterance was often identified as being stressful. Listeners may use voicing irregularity as a cue in determining whether an utterance has been generated by a person under stress.

Subject R.E.

The conventional wide-band spectrograms for this subject, shown at the top and bottom of Fig. 11, indicate a special kind of voicing irregularity. The vocal folds sometimes seem to vibrate in a mode in which every other glottal pulse is suppressed, but this mode occurs in an erratic fashion. During this mode, the individual periods of vibration exhibit a complex pattern with what appear to be several glottal pulses in one cycle. The spectrogram of "Deviation Angle" shows that the syllabic [n] is shortened and weakened. Less precise articulation is also evident in the spectrograms of "Oil Pressure"; the diphthong [ɔi] does not reach the [i] target as closely in the right-hand spectrogram as in the left-hand spectrogram. The contour spectrograms in the middle of the figure demonstrate certain differences in vowel spectra for frequencies above the second formant. For the first syllable of "Generator," the right-hand spectrogram shows less high-frequency energy. This effect does not occur in "Current," however.

Subject D.M.

The spectrograms of "Deviation Angle," shown at the top of Fig. 12, illustrate the few effects of stress that were observed for this subject. In the right-hand spectrogram, the vocal folds appear to vibrate in a mode that is sometimes called vocal fry.[3] This is a condition in which the glottal waveform is characterized by two pulses, a larger pulse and a smaller one. The spectrogram thus shows weaker vertical lines between the normal stronger lines. There is also evidence of a lack of precise articulation; no stop gap is seen for the [g] in "Angle," and there is no indication of a lowering of the velum.

Subject P.M.

The spectrograms in the middle and at the bottom of Fig. 12 illustrate some of the manifestations of stress in this subject's speech. In the spectrograms of "Generator Current," there is evidence that the subject articulates certain consonants more precisely when he is under stress. The [n] and the [t] in "Generator" are produced with a greater constriction in the right-hand spectrogram. Another feature that may be related to stress is the manner in which voicing begins following certain voiceless consonants. In the right-hand spectrogram, the onset of voicing following the [dʒ] and the [k] clearly shows glottal pulses at low frequencies, but there appears to be a delay before the pulses have high-frequency energy. It might be speculated that the subject has less precise control over voicing onset after a voiceless interval. In the spectrogram of "Oil Pressure," there appears to be less energy in the frequency region of the third and fourth formants.

Subject J.F.

A number of characteristics were observed for this subject; most of these are illustrated in Fig. 13. The voiceless consonants seem to be generally characterized by more high-frequency energy in the right-hand spectrograms. For example, in the spectrogram of "Generator Current," the [dʒ] has strong components over the entire frequency range above 2.5 KHz. In the spectrograms of all three test words, there is evidence of a less precise or slurred articulation. Examples are the [n] and [r] in "Generator," the syllabic [n] in "Deviation," and the [ng] in "Angle," There is also shortening of many of the vowels. Characteristics associated with the manner of vocal-fold vibration are also evident. Most spectrograms show less high-frequency energy during voiced intervals. Voicing irregularity and vocal fry are also noted, particularly in the spectrogram of "Deviation Angle." The rapid alternation of voiced and voiceless segments in the test word "Capsule Temperature" seems to have a profound influence on the voiced intervals; the spectrogram of this test word shows that normal vocal-fold vibration is simply not achieved.

Subject D.K.

This subject tended to use different inflection patterns in the stress and control conditions, especially for test words in the initial position. The narrow-band spectrograms shown in Fig. 14 illustrate this effect. The utterance produced in the control condition terminates with a high and slightly rising fundamental frequency, whereas the utterance produced under stress terminates with a falling fundamental frequency. The spectrogram of "Deviation Angle" shows some signs of less precise articulation; the [v] is less constricted, the [i] following the [v] has a lower third formant, and the [ə] following the [ʃ] is shortened. Both syllables in "Angle" are also shortened. There is some evidence that the glottal pulses have less energy in the frequency region of the third and fourth formants.

Subject B.E.

When this subject was stressed, he tended to speak with a more monotone voice, as is indicated in the narrow-band spectrograms shown in Fig. 15. (A similar characteristic was noted for Subject B.M. in Fig. 10.) The right-hand narrow-band spectrogram, in the vicinity of 1.6 to 1.7 sec and 1.9 to 2.0 sec, shows some noise-like energy between the individual harmonics at frequencies above 1.5 KHz. This is interpreted as evidence of vocal fry (in the case of the region 1.6 to 1.7 sec) and voicing irregularity (in the case of the region 1.9 to 2.0 sec). These characteristics are not, however, observable in the spectrogram of "Generator Current." The latter spectrogram illustrates poor articulation of the consonants [n] and [t] in "Generator."

Subject C.W.

Figure 16 shows two pairs of spectrograms for this subject, together with oscilloscope photographs of the waveform (filtered from 1.5 to 2.5 KHz) during the 0.2 sec interval just prior to the [ʃ] in "Deviation Angle." The voicing irregularity in "Deviation Angle," observable both in the spectrogram and in the waveform, has been referred to previously. The right-hand spectrograms also show evidence of less precise articulation of certain consonants and some shortening of vowels. For example, the [ℓ] in "Relative" is less constricted, the [I] in "Relative" is shortened, and both [t]'s in "Relative Velocity" are less aspirated. The [ə] in "Deviation" is all but omitted.

5. CONCLUSIONS

The addition task developed in this study appears to be a suitable vehicle for investigating the effects of stress on speech. Undoubtedly the most important feature of this task is that the experimenter can vary the difficulty of the task according to his estimate of the level of stress induced in the subject. This allows subjects with widely different tolerances for stress to be accommodated. In addition, the task is sufficiently realistic to interest and motivate most subjects. For these reasons, the task and the associated experimental procedure are recommended for use in future research.

The results of this study support the following general conclusions: There are a number of potential manifestations of stress in the acoustic speech signal. Most of the effects of stress are related to the manner in which the glottal pulses are generated in the larynx. Stress can influence the amplitude of these pulses (level), the average rate at which the pulses are generated (fundamental frequency), the contour of fundamental frequency during a breath group, the shape or frequency spectrum of each pulse, the regularity in shape of successive pulses, the initiation of glottal vibration following a voiceless interval, and the amount of turbulent noise generated along with the glottal pulses. Other effects of stress are related to articulation. The durations of phonetic segments can be altered, and the precision with which articulatory targets for vowels and for consonants are reached can be affected. The utterances produced by a given individual while he is under stress will usually exhibit only some of these effects.

For many people, the manifestations of stress are well-defined and occur consistently in most of their utterances. For other people, the manifestations of stress are more sporadic. For still other people, there are few, if any, consistent manifestations of stress.

The acoustical characteristics associated with a given effect of stress for one individual may be quite different from those for another individual. For one person, an increase in a particular variable of the speech signal is indicative of stress, and for another person, a decrease in the same variable is equally significant. An example of such a variable is the fundamental frequency. Some effects of stress, such as voicing irregularity, are even more discrepant in different individuals; while most people exhibit this characteristic only under stress, others occasionally exhibit it when they are relaxed. As long as such effects occur in a consistent manner in most of the utterances of a given individual, they may be used to predict whether other utterances by the same individual were produced under stress.

The findings of this study suggest several kinds of further work. One avenue for further research would be to examine in detail the physiological changes that underlie the manifestations of stress in the acoustic speech signal. For example, it would be of interest to determine whether the various effects of stress on the glottal output stem primarily from differences in respiratory activity (such as greater or lesser subglottal pressure), from differences in control of the laryngeal musculature, or from differences in the physiological state of the vocal folds (such as the accumulation of mucus or excessive dryness). In such an investigation, a number of physiological parameters would have to be measured during the production of speech.

Further work could also be directed toward the development and possible automation of the analysis techniques employed in this study. There are several practical applications for a device which can determine when a person is confronted with a stressful situation. One application is in the nation's space program: It may be desirable to monitor and analyze an astronaut's voice, and to use the obtained data, along with physiological and behavioral measures, for deciding whether he is able to continue his mission or whether his flight should be terminated. Although an astronaut is concerned with his personal safety, he is presumably motivated to carry his mission to completion, and hence to minimize in his verbal reports any difficulties he may encounter.

During his training in simulated spacecraft, the astronaut could serve as a subject in a task such as the one developed in this study. The test words could be phrases that he is likely to use in his communications from space. A sufficient number of voice samples could be obtained during experimental stress and control conditions to provide an unambiguous identification of the manifestations of stress in his particular speech signal. The acoustical characteristics associated with these manifestations could then be incorporated into an automatic analysis scheme which would be employed to monitor his voice during actual space travel.

6. ACKNOWLEDGMENTS

The authors wish to thank Warren M. Brodey, M.D., for his numerous contributions to this study. As a consultant, Dr. Brodey suggested the approach underlying the task and the experimental procedure. While the pilot experiments were being conducted, he recommended various improvements for the task. In addition, Dr. Brodey directed the course of the later experiments in which data were obtained for subsequent analysis.

The contributions of Dr. Vincent J. Sharkey are also appreciated. Dr. Sharkey suggested the use of the addition task.

7. REFERENCES

1. A. J. Friedhoff, M. Alpert, and R. L. Kurtzberg, "An Electro-Acoustic Analysis of the Effects of Stress on Voice." J. Neuropsychiatry 5, 266-272 (1964).

2. W. Harris, R. R. Mackie, and C. L. Wilson, "Performance Under Stress: A Review and Critique of Recent Studies." Tech. Rept. No. VI prepared under Contract Nonr 1241(00) for the Office of Naval Research, Washington, D. C. by Human Factors Research, Inc., Los Angeles, Calif. (July 1956).

3. H. Hollien, P. Moore, R. W. Wendahl, and J. F. Michel, "On the Nature of Vocal Fry." J. Speech and Hearing Res. 9, No. 2, 245-247 (1966).

4. W. B. Knowles, "Operator Loading Tasks." Human Factors 5, No. 2, 155-161 (1963).

5. R. S. Lazarus, J. Deese, and S. F. Osler, "The Effects of Psychological Stress upon Performance." Psychol. Bull. 49, No. 4, 293-317 (1952).

6. P. F. Ostwald, "Acoustic Methods in Psychiatry." Scientific American, 82-91 (March 1965).

7. J. Ruesch and A. R. Prestwood, "Anxiety: Its Initiation, Communication and Interpersonal Management." A.M.A. Archives of Neurology and Psychiatry 62, 527-550 (1949).

8. J. W. Senders, J. I. Elkind, M. C. Grignetti, and R. Smallwood, "An Investigation of the Visual Sampling Behavior of Human Observers." Rept. No. 1246, Bolt Beranek and Newman Inc., Cambridge, Mass. (May 1966).

8. APPENDIX: INSTRUCTIONS FOR SUBJECTS

First Set of Instructions

This experiment is concerned with the ability of people to read meters. Before we start the experiment, we would like you to help us check the operation of our equipment.

As you can see, in front of you are six meters. Directly below the meters is a rectangular white box which will indicate what the meters are measuring: for example, OIL PRESSURE. Your task will be to read only the middle meter in the lower row (disregard the other meters) and to say what the meter is measuring. The meter may measure anywhere from -50 (far left position) to +50 (far right position) in multiples of 5. In determining your reading, round off the value shown to the nearest multiple of 5. For example, if the meter shows a value of -34, read it as -35. Or, if the meter shows +21, read it as +20. In announcing your reading, always use the following format:

(1) Say what the meter is measuring (i.e., what is indicated by the white box)
(2) Give your reading
(3) Say again what the meter is measuring

For example, if the meter reads +31 and the white box indicates OIL PRESSURE, you will announce:

"OIL PRESSURE Plus 30 OIL PRESSURE"

Do you have any questions?

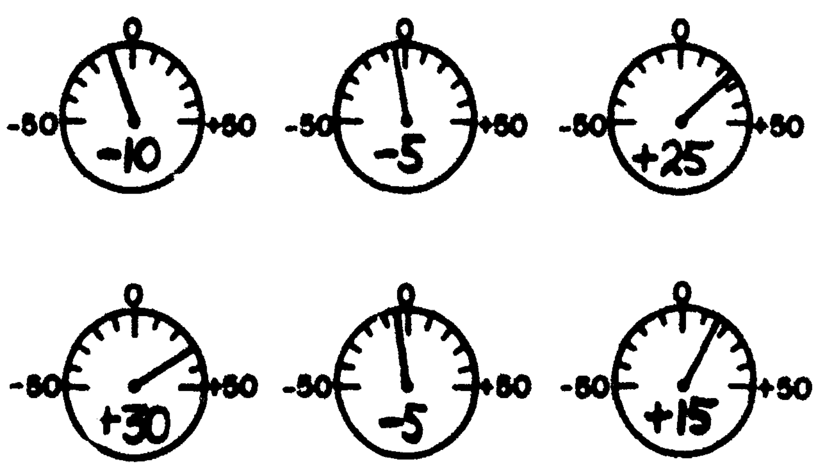
Second Set of Instructions

Pilots of aircraft and spacecraft are often required to read accurately the various meters on their instrument panels. An inaccurate reading could mean the failure of a mission and, in some instances, the death of the pilot and his crew.

We would like you to participate in an experiment which approximates the meter-reading task of pilots. Since it is difficult to create in the laboratory a realistic situation which might confront a pilot, where the misreading of a meter could have disastrous consequences, we must emphasize the importance of the task you are going to perform. To give you an idea as to the importance of your task, we will pay you 25 cents for every correct response you give. You can, depending on how well you perform your task, leave with a considerable amount of money.

In this experiment, you are to read all six meters. The meters will show values simultaneously. You are to add your six meter readings (each rounded off as before) and announce only the sum. As before, you will also announce what is being measured (i.e., what is indicated by the white box). The meters will stay on until you have made your response.

For example, if the meters display:



and the white box indicates GENERATOR CURRENT, you will announce:

"GENERATOR CURRENT Plus 50 GENERATOR CURRENT"

If your response is correct, the clear lamp at the top left of the white box will flash. If you response is incorrect, the red lamp on the right will flash.

Do you have any questions?

Third Set of Instructions

As you can imagine, a pilot must read his instruments not only accurately but also rapidly. This experiment differs from the previous one only in that the six meters will stay on for a limited time interval. You should, therefore, read and add the meter values as fast as you can, but you may announce your answer after the meters are turned off. The shorter the time interval during which the meters are on, the more you will be paid for each correct response.

You have some control over the time interval during which the meters are on. In front of you are two buttons, labeled LESS TIME and MORE TIME. If you want the meters to stay on for a shorter time so that each correct response is rewarded with more money, depress the button marked LESS TIME. On the other hand, if most of your responses are incorrect and you want the meters to stay on longer, depress the button marked MORE TIME.

Use the same format as before in announcing your response.

Do you have any questions?

Fourth Set of Instructions

The experiments in which we measured the accuracy and speed of your responses are now over. Before you leave, however, we would like you to again read one meter, as in the first experiment. You will read only the middle meter in the lower row (disregard the other meters) and say what the meter is measuring. The meter will stay on until you have made your response.

Use the same format as before in announcing your response.

Do you have any questions?

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Bolt Beranek and Newman, Inc. 50 Moulton Street Cambridge, Massachusetts 02138	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP N/A

3. REPORT TITLE

THE EFFECTS OF TASK-INDUCED STRESS ON SPEECH

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Scientific. Final. 1 July 1966 - 30 June 1967 Approved 14 Sept.'67

5. AUTHOR(S) (First name, middle initial, last name)
Michael H. L. Hecker Gottfried von Bismarck
Kenneth N. Stevens Carl E. Williams

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
25 August 1967	71	8

8a. CONTRACT OR GRANT NO. AF 19(628)-6052 b. Project, Task, Work Unit Nos. 5628 02 01 c. DoD Element 6144501F d. DoD Subelement 681305	9a. ORIGINATOR'S REPORT NUMBER(S) BBN Report No. 1542 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) AFCRL-67-0499

10. DISTRIBUTION STATEMENT
Distribution of this document is unlimited. It may be released to the Clearinghouse, Department of Commerce, for sale to the general public.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
Tech, Other	Air Force Cambridge Research Laboratories (CRB) L. G. Hanscom Field Bedford, Massachusetts 01730

13. ABSTRACT

In order to induce stress in an experimental subject, a task involving the addition of numbers under time pressure was developed. The subject was required to read six meters and to announce the sum of his readings, together with a test word. By controlling the duration of the meter display, the experimenter could vary the level of stress induced in the subject. For each of ten subjects, numerous verbal responses were obtained while the subject was under stress and while he was relaxed.

Contrasting responses containing the same test word were assembled into paired-comparison listening tests. Listeners could identify the stressful responses of some subjects with better than 90 percent accuracy and of others only at chance level. The test words from contrasting responses were analyzed with respect to level and fundamental frequency, and spectrograms of these test words were examined. The results indicate that stress can produce a number of characteristic changes in the acoustic speech signal. Most of these changes are attributable to modifications in the amplitude, frequency, and detailed waveform of the glottal pulses. Other changes result from differences in articulation. Although the manifestations of stress varied considerably from subject to subject, the test words of most subjects exhibited some consistent effects.

DD FORM 1 NOV 65 1473 (PAGE 1)
S/N 0101-807-6811

Unclassified
Security Classification

A-31408

Unclassified

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Non-Verbal Contents of Speech						
Physiological Stress						
Psychological Stress						
Speech Analysis						
Speech Communication						
Speech Production						
Task-Induced Stress						